frontiers | Research Topics

Credition - an interdisciplinary approach to the nature of beliefs and believing

Edited by
Rüdiger J. Seitz, Hans Ferdinand Angel, Raymond Paloutzian and Ann Taves

Published in
Frontiers in Psychology
Frontiers in Behavioral Neuroscience
Frontiers in Pain Research

FRONTIERS EBOOK COPYRIGHT STATEMENT

The copyright in the text of individual articles in this ebook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this ebook is the property of Frontiers.

Each article within this ebook, and the ebook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this ebook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or ebook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714
ISBN 978-2-8325-2993-5
DOI 10.3389/978-2-8325-2993-5

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Credition - an interdisciplinary approach to the nature of beliefs and believing

Topic editors

Rüdiger J. Seitz — Heinrich Heine University of Düsseldorf, Germany
Hans Ferdinand Angel — University of Graz, Austria
Raymond Paloutzian — Westmont College, United States
Ann Taves — University of California, Santa Barbara, United States

Citation

Seitz, R. J., Angel, H. F., Paloutzian, R., Taves, A., eds. (2023). *Credition - an interdisciplinary approach to the nature of beliefs and believing*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-2993-5

Intro image

Image used for the International Credition Conference in Hannover, 20 - 22 October 2021.
Transparent head showing the brain (© ADIMAS - FOTOLIA 45501197/adimas/Fotolia) surrounded by binary code symbolizing ambient information'

This RT has been partially funded by grants obtained by Professor Rüdiger Seitz via the Volkswagen Foundation, Siemens Healthineers, and the Betz Foundation.

Table of contents

Introduction

Belief formation and updating

Neural processes underlying believing

Beliefs as conceptual expressions

Abnormalities of believing

Believing enables decisions

Believing and social life

Believing and religion

frontiers | Frontiers in Behavioral Neuroscience

TYPE Editorial
PUBLISHED 23 May 2023
DOI 10.3389/fnbeh.2023.1217648

Check for updates

OPEN ACCESS

EDITED AND REVIEWED BY
Richard G. Hunter,
University of Massachusetts Boston,
United States

*CORRESPONDENCE
Rüdiger J. Seitz
✉ seitz@neurologie.uni-duesseldorf.de

RECEIVED 05 May 2023
ACCEPTED 10 May 2023
PUBLISHED 23 May 2023

CITATION
Seitz RJ, Angel H-F, Paloutzian RF and Taves A (2023) Editorial: Credition—An interdisciplinary approach to the nature of beliefs and believing. *Front. Behav. Neurosci.* 17:1217648. doi: 10.3389/fnbeh.2023.1217648

COPYRIGHT
© 2023 Seitz, Angel, Paloutzian and Taves. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Editorial: Credition—An interdisciplinary approach to the nature of beliefs and believing

Rüdiger J. Seitz[1]*, Hans-Ferdinand Angel[2], Raymond F. Paloutzian[3] and Ann Taves[4]

[1]Department of Neurology, Medical Faculty, Centre of Neurology and Neuropsychiatry, LVR-Klinikum Düsseldorf, Heinrich-Heine-University Düsseldorf, Düsseldorf, Germany, [2]Institute of Catechetics and Religious Education, Karl Franzens University Graz, Graz, Austria, [3]Department of Psychology, Westmont College, Santa Barbara, CA, United States, [4]Department of Religious Studies, University of California, Santa Barbara, Santa Barbara, CA, United States

KEYWORDS

beliefs, social interaction, rituals, perception, religion, trust, appraisal, decision making

Editorial on the Research Topic
Credition—An interdisciplinary approach to the nature of beliefs and believing

Introduction

On the 20th through the 22nd of October 2021, the international symposium on "Creditions—An interdisciplinary Challenge" took place in the Conference Center of the Volkswagen Foundation in Schloss Herrenhausen, in Hannover, Germany. Due to the Sars2-Covid-19 pandemic the symposium had a hybrid format which allowed the participation of those unable to attend in person. Our aim from the outset was to publish a book based on the symposium presentations. Thus, we are delighted to introduce this e-book consisting of 42 chapters in total on the topic of belief and believing. We are grateful to the Volkswagen Foundation, Hannover, Germany, Frontiers Publishers, Lausanne, Switzerland, Siemens Healthineers, Erlangen, Germany, and the Anton-Betz-Foundation of the Rheinische Post e.V., Düsseldorf, Germany, for their generous support of this conference and publication of its proceedings.

The start of the credition project dates back to the first meeting under the auspices of the Karl Franzens University in Graz, Austria, in 2011 (https://credition.uni-graz.at/de/credition-research/). Our point of departure was the hypothesis that information processing in the human brain concerning external events in the environment and subjective internal states affords believing and predictive control of behavior (Seitz and Angel, 2012). Thus, this project focused on how to assess subjective experience with objective measures as discussed recently by Pauen and Haynes (2021). We hypothesized that the neural processes underlying believing constitute a domain similar to those for cognition and emotion, and therefore advanced the neologistic term "credition" to represent this domain (Angel et al., 2017). The plural form, creditions, is an umbrella term that signifies the neural subfunctions that constitute the category of processes of believing, as are similarly present in the categories of emotions, perceptions, and actions (Angel). The credition concept concurs well with the notion that internal states encode beliefs about the external world that involve belief formation, belief updating, and the transmission of beliefs to others, which yields shared beliefs (Albarracin and Pitliya).

The present chapter puts the contributions of this e-book into perspective in light of the current interdisciplinary research on belief and believing.

Belief formation and updating

From the perspective of cognitive science, belief formation and updating result from the neuropsychological processes that afford believing. These processes consist of perception of external information, valuation in terms of subjective relevance or meaning, predictive coding of subsequent behavior, and encoding of this composite information in and retrieving it from memory (Seitz et al., 2018; Seitz and Angel, 2023). Perception involves unconscious multisensory integration within and across modalities and potentially conscious awareness of percepts that result (Firestone and Scholl, 2016). Because far more information is in the environment than any organism can perceive and process, organisms have evolved to perceive and process the information they need to survive. Humans and some other animals do so by segmenting information by a cognitive process that divides the information into chunks with a beginning and an end (Taves and Paloutzian). Furthermore, the processes of believing function to stabilize a given perception out of the myriad rapidly changing external stimuli in terms of personal meaning, which allows for socially adaptive behavior. Global meaning encompasses foundational beliefs, values and goals, and a subjective sense of meaningfulness, whereas situational meaning entails the appraisal of an experience (Park). Interestingly, the gut-brain-gut communication network is part of the interoceptive circuits that enable a person to sense and interpret the physiological condition of the body and regulate its autonomic and mental activity (Holzer). Ultimately, the processes of believing constrain an individual's behavior in a stochastically predictable way (Seitz et al., 2018). As humans trust their beliefs, the beliefs provide a temporarily reliable link between a person's past experience and his/her future behavior.

With respect to the bottom-up processing of external information that can become the object of believing, three categories of beliefs have been differentiated: (a) empirical beliefs about objects and facts, (b) relational beliefs about events including human interactions, and (c) conceptual beliefs about narratives including those held in societies (Seitz and Angel, 2020). Empirical and relational beliefs occur below the level of awareness, and thus correspond to so-called primal beliefs as depicted in a multi-level scheme (Table 1). In contrast, conceptual beliefs are mediated by language and are objects of conscious awareness. This conceptualization accounts for the hierarchically nested structure of the three-levels of believing processes, i.e., the physical level, the interpersonal level, and the social level (Sugiura et al., 2015). In essence, the ability to believe expands human cognitive, sensory, and perceptual dynamics and is essential for the human ability to engage with and shape the world, as is evident from phylogenetic evolution (Fuentes). This accords with van Eyghen' claim that explaining belief and believing from an ontogenetic perspective is more parsimonious than from a phylogenetic account, because there is no need to postulate anything beyond the plasticity of the human brain and mind.

TABLE 1 Beliefs as the result of the processes of believing.

Input level	Objects	Events	Narratives
First person level	I believe that	I believe him/her	I believe in
Third person level	Empirical beliefs	Relational beliefs	Conceptual beliefs
Meta analytic level	Primal beliefs		Autobiographic Religious Political

For comparison, placebo effects rely on the brain's ability to integrate contextual information in the environment with prior experiences, and are likely due to emotional re-appraisal strategies and cognitive-evaluative processes. Only very strong placebo interventions, such as those induced by classical conditioning, may affect early sensory processes in a significant manner (Meissner). Upon believing, we assume that what we believe is true and that it correctly reflects the environment. Thus, what we personally believe is true may be mistaken, e.g., by visual plausibility (Adelson, 1993) or by confused timing of thought (Bear et al., 2017). For reasons of space, we do not enter the discussion of the different concepts of truth (McLeod, 2021) nor do we refer to different functions of truth, as for instance in the debate about "narratives" (Mercier, 2020). Rather, we refer to the empirical evidence that truth judgments are based on a bias to judge incoming information from the environment as true, so long as there is an ease of processing whether assertions match information stored in memory (Brashier and Marsh, 2020). Sensory perceptions are typically processed with ease and thereby construct our experience of our environment. This is no different from how language-based information is assembled. Typically, repetition of statements that facilitate the subjective ease of processing has been shown to increase the likelihood that a statement is judged true (Wang et al., 2016). Nevertheless, language comprehension can be difficult and may even require a third person's interpretation. Usually, what a person says or does is taken to reflect what he/she is believing. In fact, belief congruity, social congruence, and message repetitions have been proposed to enhance the probability that implausible and false information may be accepted as true (Levine, 2022). However, deceptive intentions by other persons have to be taken into consideration as well. Accounting for these different aspects, Connors and Halligan have proposed a five-stage model for belief formation that involves a triggering sensory precursor, meaning attribution, belief evaluation, belief acceptance, and effects of beliefs.

In accordance with these observations we claim that humans, like non-human primates, are engaged in believing that their senses provide a true image of their environment. Although there are reasons to decide amongst alternatives about how to behave, to believe is a mandatory function which enables a subject to develop preferences to regulate behavior in an ecologically adequate fashion in a complex environment. However, there are non-evidential reasons, be they embedded in religions, worldviews, or secular ideologies, for believing (Longheed and Simpson, 2017). Similarly,

the notion that some of our beliefs are under our control—we manage the cognitive mechanisms that issue them and control whether they operate in the right environment (Visala)—makes it likely that we will underestimate the fluidity of beliefs brought about by new information from the environment (Seitz et al., 2018). This fluidity needs to be differentiated from the colloquial saying, "There is good reason to believe that...," because it is a meta-cognitive statement from a third person perspective. Such a statement conveys that the person who is stating it judges the thing in question to be similar to how it is judged by the person whose behavior he/she observed.

Neural processes underlying believing

The neural representations involved in the formation and updating of primal beliefs about objects and beliefs are pre-linguistic in nature and are maintained in large-scale cortico-subcortical networks in the human brain (Seitz). The cortical structures involved include the dorsolateral prefrontal cortex, the parietal cortex, and the so-called pre-supplementary motor area in the dorsomedial frontal cortex. When people believe that they have recognized a target, this network including subcortical structures like the basal ganglia, thalamus, and amygdala become active as was shown in a functional magnetic resonance imaging (fMRI) study when subjects were asked to indicate when they recognized emotions in slowly evolving facial stimuli (Sonnberger et al.). Of particular relevance are the brain structures that are part of the affect regulating system. For example, in another fMRI study it was found that the belief that a leader is transformational triggers neural activations in the follower's reward circuitry that correlate with the follower's level of motivation (Bergner et al.). Most recently, a large fMRI study on more than 900 volunteers has shown that emotions can enhance memory encoding of pictures which is mediated by a large circuit of interconnected brain areas including cortical areas, the hippocampal formation, the amygdala as well as the thalamus and cerebellum (Fastenrath et al., 2022). Moreover, when subjects were required to listen to stories, fMRI revealed an enhanced activity in the widespread cortical semantic system related to specific semantic domains or groups of related concepts (Huth et al., 2016). Conversely, transient inactivation of these areas, the left inferior frontal gyrus, by transcranial magnetic stimulation was found to reverse the habitual tendency to discount bad news in belief formation (Sharot et al., 2012).

A meta-analytic research project revealed that mindfulness can be acquired by meditation techniques and lead to emotional regulation, and to monitoring perception and behavior with particular emphasis on increasing the experiential phenomenological self and reducing self-relational thoughts of the narrative self (Weder). Self-referential thinking during mindfulness and self-relational thinking in the narrative self relies on the default mode network including the dorsal and medial prefrontal cortex, and posterior cingulate cortex. These findings correspond well to the notion that self-estimates of abilities like self-esteem, self-concept, and self-efficacy are conceptually close to beliefs (Neubauer and Hofer). Furthermore, it has been suggested that the common cognitive bias underlying the multidimensionality of self-transcendence is related to a sense of self-agency, indicating the possibility that the bias is caused by a process that controls the neural networks involved in multilevel forward model prediction (Sugiura). From a phylogenetic point of view it is noteworthy that when monkeys viewed other monkeys, a number of processes took place. They included the recall of novelty and emotional significance from memory of previous experiences with other macaques, the novelty of the individual seen in a mirror, innate fear, etc. (Bretas et al.). Specifically, the belief that the macaque in the mirror is a reflection of the self was found to be expressed in the form of mirror self-recognition behavior.

Social interactions of individuals rely on believing the bodily and verbal expressions of the counterpart, which can be suspected to involve empathy. Using a new model of empathic learning using a feedback loop it was found that changes in inter-brain coupling in the inferior frontal gyrus represent a core component of affect empathic reactions (Shamay-Tsoory). Moreover, an embodied approach to abstract words and cognitive concepts may shed light onto the process of building and revising beliefs, specifically suggesting that beliefs, much like other conceptual domains, can be grounded in actual experiences and their complexity (Buccino and Colagè). Furthermore, brain imaging results in healthy volunteers of Caucasian and Chinese ethnicity suggest that the development of culturally specific beliefs is brought about by culture–brain interactions via the practice of behaviors and by direct culture–brain interactions that are based on distinctive neurocognitive processes (Han et al.).

Beliefs as conceptual expressions

Only a small proportion of information enters someone's conscious awareness and can then be expressed from first-person perspective as "I believe …" (Oakley and Halligan, 2017; Seitz and Angel, 2020). Such a proposition is a probability statement that signifies by means of verbal behavior an affective involvement of the speaker. It is used with a slightly different phrasing for objects, events and narratives as summarized in Table 1. These statements are different from a confidence statement (Ülkümen et al., 2016). Accordingly, people use the verb "believe" in a highly differentiated fashion and in different contexts compared to how they use the verb "think." Empirical evidence suggests that people use "believe" preferentially in religious contexts, whereas they say "think" when they refer to a confidence statement about facts (Heiphetz et al., 2021). In contrast, it is uncommon to use the noun belief for such a statement (It is my belief that …), although it is a common expression from the third-person perspective (It is his/her belief that …). Typically, the content of such a belief is specified in certain areas of discourse such as religion, morality, politics, etc. Although commonly done in English, one should be aware that labeling is a post-hoc attribution from a meta-analytic perspective (Seitz et al., 2022). Thus, the belief in question is brought about by inferential thinking of an observer and attributed to a behavioral outcome such as a verbal statement or an action. Accordingly, the labels *political*, *religious*, *moral*, and *social* belief involve the tacit claim that believing can be classified from a third-person top-down perspective according to putative epistemological entities, such as religion, politics etc., (Table 1). In fact, these entities are language-based narratives that represent what we have called conceptual

beliefs (Seitz and Angel, 2020). Probably related to a teleological view, the specificity of such conceptual contents of beliefs has been questioned (Oviedo and Szocik, 2020). Furthermore, a *post-hoc* attribution is hardly compatible with a general neuroscientifically grounded model of belief formation and updating, as realized in parallel organized cortico-subcortical networks affording predictive processing (Friston et al., 2017; Seitz et al., 2018).

It is important to realize that there are intriguing linguistic issues concerning the notion of beliefs and believing (Angel). In English one can speak about beliefs in plural. In contrast, in German the term for belief appears monolithic, as it does not have a plural form. When one has to translate beliefs (plural) into German, most likely instead of belief (=GLAUBEN) another term (MEINUNG = opinion) will be used because it can occur in both the singular and the plural form. Yet, one has to acknowledge that "opinion" lacks an affective meaning, in contrast to "belief." Thus, texts that were translated from English into German may suffer the lack of linguistic clarity. For instance, one can have a religious GLAUBEN but religious MEINUNG does not make sense. Furthermore, in German there is no equivalent term for believing. In contrast, the phrase "processes of believing" can be translated into German. Therefore, we have to acknowledge that how language is used indicates cognitive assumptions about prior knowledge that is likely to influence the adoption of new information and conclusions (Madzarevic). Nevertheless, credition was said to afford openness of the self to the freedom and play that are fundamental to being human (Davies).

Notably, there is a tight link between belief and knowledge, as knowledge has traditionally been defined in philosophy as justified true belief. It is important to note, however, that Popper replaced the problem of justification with the issue of criticism, which is an argument for a fluid character of beliefs (Diller, 2006). Nevertheless, from a philosophical point of view beliefs have been dichotomized into categorical (yes/no) beliefs and graded beliefs. While the former are logically coherent and deductively closed, the latter are probabilistically coherent with a probability of <1.0 (Dietrich). Likewise, doxastic logics lead to propositions concerning beliefs ("it is believed that"), whereas deontic logics result in prescriptions ("it is obligatory that"). The interesting question is how these beliefs can be revised Vestrucci). Beliefs, however, may also reflect a property of the believing person. For example, according to the concept of representationalism, a given representation with the content P may be deployed in reasoning. For comparison, according to dispositionalism, a person may believe a given proposition, because she/he is disposed to act and react in this way (Schwitzgebel). Furthermore, the belief that a person is epistemically confident about something is likely to be formed and revised differently from a belief that is central to a person's identity or heart (van Leeuwen). Nevertheless, one should be aware that these discussions deal with *post-hoc* theoretical reasoning but not with cognitive science of belief formation and updating.

Abnormalities of believing

Diseases of the brain may disrupt any of the processes of belief formation and updating, as for example in the alien limb syndrome, agnosia, hallucinations, and delusions (Seitz, 2022). For example, empirical studies have shown that in altered sensorimotor processing, self-monitoring can link hallucinations of presences to the detection of human agents (Vehar et al.). From a pathophysiological perspective it is noteworthy that brain lesions affecting the dorsolateral and ventromedial prefrontal cortex as well as the posterior superior temporal cortex were found to facilitate the occurrence of religious beliefs, mystical experience, and ideological commitments (Cristofori et al.).

Furthermore, after traumatic experiences people have been shown to make meaning to reduce discrepancies between situational and global meanings, with a greater reduction in the size of discrepancies predicting better adjustment following trauma (Park). Similarly, in the Covid-19 pandemic, patients with affective disorder were more uncertain and experienced fewer positive emotions than healthy controls, although both groups did not differ in vaccination status (Dalkner et al.). Particularly, in psychotic disorders and a wide range of other neuropsychiatric conditions abnormalities of belief formation may result in discrepancies between bodily expressions and verbal reports. Such discrepancies may cause distrust in the addressee(s) and eventually may destroy social bonds. However, because beliefs are subject to change, people may adapt their behavior and can create new experiences—often during social interactions—which may help them to leave abnormal beliefs behind (Pott and Schilbach) and facilitate the speculation that psychotherapeutic interventions may become operative via socio-verbal interaction.

Believing enables decisions

Decision making has been the object of scientific research for many years opening broad perspectives in the theoretical and practical areas of the sciences. The neural processes affording decision making have been studied in animals including non-human primates and mammals as well as in humans using neurophysiological and neuroimaging techniques. Moreover, the roles of attention, perception and choice-consistency have been explored recently (Nitsch and Kalenscher, 2021). Owing to the notion that meaning making and affective relevance are inherent in the processes of belief formation, a tight link between believing and the establishment of preferences can be postulated. Preferences allow for predictive coding and, thus, are key factors in decision making and selection of behavior. Empirical findings support the notion that our preferences evolve endogenously during the process of making decisions between equally preferred items (Voigt et al.). Therefore, self-determined, subjective cognitive concepts, such as our preferences, might be emergent consequences of the particulars of the decision scenario itself. Findings from functional neuroimaging studies support the view that the orbitofrontal cortex contributes to expectation-guided decision-making by enabling us to simulate the consequences of our choices (Kahnt). Moreover, it was found in choice tasks that value of the items and confidence in the decision involve large parts of the medial prefrontal cortex with a specific activation for value in the ventral portion and for confidence more dorsally in the anterior portion (Claris and Pessiglione, 2022).

Against this background and with respect to many failed engineering projects, it ought to be questioned whether engineers usually make rational decisions during product development. How to support decision-making is therefore a central topic in complex decision situations (Kranabitl and Faustmann) including economic decisions. For such purposes, an elaboration of artificial intelligence modeling of the capacity to reflect, rationalize, and communicate has been developed to support and even improve decision making (Lumbreras).

Believing and social life

Individuals are members of social groups. After birth these groups are families that used to belong to tribes and, nowadays, typically are inhabitants of a village or town. Given these contexts, a person's behavioral decisions can be expected to evoke reactions in those to whom they are addressed or in other group members who are bystanders. To review the wealth of historical, philosophical, anthropological and psychological literature on this issue would far exceed the limits of this article. Nevertheless, in what follows we highlight some aspects of creditions that are pertinent to the relationships between believing and social life.

It is well-known that people can communicate the content of their beliefs as personal statements and can repeat the statements of others to themselves or other people. The power of language is that we can express our thoughts and emotions verbally, although we need to accept that in describing emotions and thoughts we are limited by the words we use (Abukhalaf, 2021). Linguistic research has shown that beliefs are based on the reliability and solidity of our knowledge and are typically described by abstract rather than concrete concepts (Borghi et al.). Thus, verbal expressions enable us to begin to understand the conceptual beliefs of other people. Above that, the exchange of verbal information typically benefits from the consistency between a person's verbal statements and his/her bodily expressions, because the person then appears particularly trustworthy (Seitz). Importantly, the transmission of narratives among members of a group can lay the groundwork for social cooperation within and possibly between groups. Whereas reputation has been found to sustain cooperative relationships among unrelated individuals in social groups and systems (Romano et al., 2021), another key to promoting prosocial behavior within a group may be reciprocation among the group members (Teehan, 2006). None the less, morality comes into play here because it promotes within-group cohesiveness and empowers individuals to protect their offspring (Teehan, 2006). Interestingly, the expectation to behave morally is not necessarily extended to individuals outside one's own group.

Consequently, it has been proposed that believing includes a component of trust that can be expressed in verbal communications, including those that convey information beyond one's personal experience. This degree of acceptance or trust probably also applies to news as well as to norms and promises within social groups. Granting trust may thereby be considered as a basis for social cooperation and group cohesion. It has been shown that the assignment of trust is learned by employing predictive coding, as is manifest in the processes of believing (Seitz). At the neural level, learning about the assignment of trust has been shown to involve the medial frontal cortex for confirmatory evidence of trust and to involve the lateral prefrontal cortex for alternative, untrustworthy outcomes (Akaishi et al., 2016). Thus the processes of believing are important neural functions that may ultimately be the springboard for the evolution of human social life and the development of culture and civilization (Fuentes). Thus, although believing is essential for creating preferences that afford behavioral decision making (see above), these processes are continuously modified by confirmatory or contradictory information (see above). Accordingly, moral and social beliefs are not stable entities that change only when there is dissonance between them (Dalege and van der Does, 2022). This view casts doubt on the notion that moral beliefs are necessarily explicit conceptual *post-hoc* descriptions (see above) and that beliefs about social networks are by definition implicit and formed in a pre-linguistic fashion (Korman et al., 2015; Seitz and Angel, 2020). Consequently, these different types of beliefs have to be assumed to compete within conscious and unconscious awareness. This raises the interesting question of whether the stability (or changeability) of beliefs is due more to external information or the individual's affect, particularly in different times or contexts.

Believing and religion

In recent years, there has been an increasing interest in the multilevel interdisciplinary research on the cultural evolution of religion and spirituality (Paloutzian and Park, 2013; Feierman and Oviedo, 2020). Originally, the Ancient Greek terms for *belief* and *to believe,* namely *πίστις (pístis)* and *πιστεύειν (pistéuein),* were not exclusively related to religious experiences. But from about the 4th century onwards the expansion of Christianity linked beliefs more and more with *Christian beliefs*. Thereafter, since the end of the Middle Ages, the notion *religious beliefs* became common. In consequence, the term and the concept(s) of *religion* became predominant as a conceptual framework for understanding religious experiences. Regardless that the understanding of religion changed profoundly during history, the development of the concept of religiosity was virtually neglected (Angel, 2022). From the cognitive science perspective, however, religiosity, not religion, is the relevant focus for understanding religious experiences (Angel, 2020). More recently, the evolutionary and cognitive accounts of religious beliefs have challenged the justification for believing in religious propositions (Teehan, 2014). Justified religious beliefs have been defined as beliefs that are consistent with the beliefs and grounds of belief employed in a given belief tradition (Teehan, 2014). It is widely accepted that religious beliefs exert a profound social impact. On the individual level they have the pragmatic aspect that they allow persons to make sense of their lives and of the world they live in. On the social level, they are said to promote inter-individual cooperation and to regulate inter-group conflict and competition within ethnic groups (Norenzayan et al., 2016). This is probably enhanced by ritualistic synchrony in religious acts that has been found to play a key role in cultural evolution (Gelfand et al., 2020). In correspondence with this notion, Geertz (2013) proposed that the co-evolution of genes and culture is a mover of the cultural evolution of religion.

It has been suggested that religious beliefs are brought about by a number of deeply engrained psychic functions such as agency detection, mentalizing, or dualistic reasoning (van Elk).

People seemingly tend to attribute significance to information from sources they deem trustworthy. Specifically, empirical evidence points toward the role of cultural scaffolding and explicit teaching for endorsing supernatural beliefs (van Elk). Furthermore, in empirical studies on over 2,000 participants from different religious traditions in the United States, Ghana, Thailand, Vanuata and China, it was found that the power of the cultures in combination with individual differences shapes what feels real to the senses such as gods and spirits (Luhrman et al., 2021). Also, based on the exploration of classical Buddhist theories, Jed Forman argues that higher-order cognitive processes, like reflection on beliefs, may not only manipulate how we see our environment but also may generate a platform for what we see. Consistent with this notion, it has been proposed that various ways to purify the mind and develop its potential can be found in ancient Buddhist sutras (Du). It is noteworthy that Islamic thought contains opinions, positions, and sayings that have been transcended in many respects to keep pace with the current questions and developments posed by the socio-cultural environment of contemporary Muslims. Therefore, the Editors regret that of the two papers addressing the Islamic background one was withdrawn by the authors while the other did not pass the publication process. Even so, we hypothesize that the fundamental logic that underpins the processes of believing applies not only to everyday life, but also to religions including Islam.

Conversely, in Western countries the increasing number of non-religious people are moving away from the traditional religious narratives that provided meaning and structure around the dead body for both themselves and others (Applewhite). Consequently, they also introduce new kinds of meaning that are likely to affect values and beliefs around environmentalism, secularism, economics as well as traditions outside of religion. Observations of this sort may raise questions about the decreasing appeal of the promises that are central in traditional religious belief systems but similarly also in political ideologies.

Conclusions

This e-book provides an up-to-date overview of how the introspective experience of believing something can be an issue of cognitive science and philosophy (Pauen and Haynes, 2021). On the behavioral or phenomenological level we have summarized the accumulating evidence suggesting that believing involves bottom-up processes that empower humans to select their behavior according to implicit as well as explicit, e.g., verbally coded, preferences. Also, we have described that the resulting beliefs typically are labeled semantically from a *post-hoc*, third-person perspective based on top-down inferential thinking. On the neural level the processes of believing were shown to be implemented in large-scale brain circuits. Whether functional imaging can show neural processes or representations such as social event knowledge or beliefs is an issue of a long discussion (Krueger et al., 2009). This e-book does not pertain only to the biological sciences but also to the theoretical sciences and the humanities. We hope that it can stimulate empirical and theoretical work to elucidate the driving forces of how humans have shaped their civilizations as well as the foundations of art.

Author contributions

RJS: writing. H-FA, RFP, and AT: editing. All authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Abukhalaf, A. H. I. (2021). Exploring emotional intelligence from a linguistic perspective. *Acad. Lett.* 3, 3574. doi.org/10, 20935./AL3574

Adelson, E. H. (1993). Perceptual organization and the judgment of brightness. *Science* 262, 2042–2044. doi: 10.1126/science.8266102

Akaishi, R., Kolling, N., Brown, J. W., and Rushworth, M. (2016). Neural mechanisms of credit assignment in a multiclue environment. *J. Neurosci.* 36, 1096–1112. doi: 10.1523/JNEUROSCI.3159-15.2016

Angel, H. F. (2020). "A History of the evolution of religion: from religion to religiosity to the processes of believing," in *The Evolution of Religion: How Biology, Psychology, and Theology Interact*, eds Feierman J, Oviedo L. Routledge/Taylor and Francis, New York. pp 87-103

Angel, H. F. (2022). *Credition. Fluides Glauben.* Kultur- und Wissenschaftsgeschichte von einem blinden Fleck und seinem Ende. Deutscher Wissenschaftsverlag, Baden-Baden, Germany, pp 406-442

Angel, H. F., Oviedo, L., Paloutzian, R. F., Runehov, A. L. C., and Seitz, R. J. (Eds.). (2017). *Processes of Believing: The Acquisition, Maintenance, and Change in Creditions.* Heidelberg: Springer. doi: 10.1007./978-3-319-50924-2

Bear, A., Fortgang, R. G., Bronstein, M. V., and Cannon, T. D. (2017). Mistiming of thought and perception predicts delusionality. *Proc. Natl. Acad. Sci. USA.* 114, 10791–10796. doi: 10.1073/pnas.1711383114

Brashier, N. M., and Marsh, E. J. (2020). Judging Truth. *Ann. Rev. Psychol.* 71, 499–515. doi: 10.1146/annurev-psych-010419-050807

Claris, N., and Pessiglione, M. (2022). Values, confidence, deliberation: a functional partition of the medial prefrontal cortex demonstrated across rating and choice tasks. *J. Neurosci.* 42, 5580–5592. doi: 10.1523/JNEUROSCI.1795-21.2022

Dalege, J., and van der Does, T. (2022). Using a cognitive network model of moral and social beliefs to explain belief change. *Sci.Adv.* 8, eabm0137. doi: 10.1126./sciadv.abm0137

Diller, A. (2006). "Constructing a comprehensively anti-justificationist position," in *Karl Popper: A Centenary Assessment*, eds Jarvie I, Milford K, Miller D. Vol II, Chapter 28, (Ashgate, London) pp 119-129

Fastenrath, M., Spalek, K., Coynel, D., Loos, E., Milnik, A., Egli, T., et al. D J-F. (2022). Human cerebellum and corticocerebellar connections involved in emotional memory enhancement. *Proc. Natl. Acad. Sci. USA* 119, e2204900119. doi: 10.1073/pnas.2204900119

Feierman, J., and Oviedo, L. (Eds.). (2020). *The Evolution of Religion: How Biology, Psychology, and Theology Interact*. New York: Routledge/Taylor and Francis

Firestone, C., and Scholl, B. J. (2016). Cognition does not affect perception: evaluating the evidence for "top-down" effects. *Behav. Brain Sci.* 39, e229. doi: 10.1017/S0140525X15000965

Friston, K. J., Parr, T., and Vries, d. e. B. (2017). The graphical brain: belief propagation and active inference. *Network Neurosci.* 1, 381–414. doi: 10.1162/netn_a_00018

Geertz, A. W. (2013). The meaningful brain. Clifford Geertz and the cognitive science of culture. In: Xygalatas D, McCorkle WW Jr.. (eds). Mental culture. Classical social theory and the cognitive science of religion. Acumen, pp 176-254

Gelfand, M. J., Caluori, N., Jackson, J. C., and Taylor, M. K. (2020). The cultural evolutionary trade off of ritualistic syndrony. Philos Trans R Soc B 375, 20190432. doi: 10.1098/rstb.2019.0432

Heiphetz, L., Landers, C. L., and van Leeuwen, N. (2021). Does "think" mean the same thing as "believe"? Linguistic insights into regliogious cognition. Psychol Religion Spirituality 13, 287–297. doi: 10.1037/rel0000238

Huth, A. G., de Heer, W. A., Griffiths, T. L., Theunissen, F. E., and Gallant, J. L. (2016). Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature* 532, 453–458. doi: 10.1038/nature17637

Korman, J., Voiklis, J., and Malle, B. F. (2015). The social life of cognition. *Cognition* 135: 30-35. doi: 10.1016/j.cognition.11, 005.

Krueger, F., Barbey, A. K., and Grafman, J. (2009). (2009). The medial prefrontal cortex mediates social event knowledge. *Trends Cogn. Sci.* 13: 103-109. doi: 10.1016/j.tins.09, 001.

Levine, T. R. (2022). Truth-default theory and the psychology of lying and deception detection. Curr Opin Psychol 47, 101380. doi: 10.1016/J.copsyc.2022.101380

Longheed, K., and Simpson, R. M. (2017). Indirect epistemic reasons and religious belief. *Religious Studies* 53, 151–169. doi: 10.1017/S0034412516000202

Luhrman, T. M., Weisman, K., Aulino, F., Brahinsky, J. D., Dulin, J. C., Dzokoto, V. A., et al. (2021). Sensing the presence of gods and spirits across cultures and faiths. Proc Natl Acad Sci USA 118, e2016649118. doi: 10.1073/pnas.2016649118

McLeod (2021). Trust. In Zalta EN. (ed). The Stanford Encyclopedia of Philosophy. Stanford University, Stanford, USA

Mercier, H. (2020). Not born yesterday: The science of who we trust and what we believe. Princeton University Press, Princeton, USA

Nitsch, F. J., and Kalenscher, T. (2021). Influence of memory processes on choice consistency. R Soc Open Sci 8, 200308. doi: 10.1098/rsos.200308

Norenzayan, A., Sharif, A., Gervin, W. M., Willard, A. K., McNamara, R. A., Slingerland, E., et al. (2016). The cultural evolution of prosocial religions. Behav Brain Res 39, e1. doi: 10.1017/50140525X14001356

Oakley, D. A., and Halligan, P. W. (2017). Chasing the rainbow: the non-conscious nature of being. Front Psychol 8, 1924. doi: 10.3389/fpsyg.2017.01924

Oviedo, L., and Szocik, K. (2020). Religious and other beliefs: how much specificity? Sage Open, 1-11; doi: 10.1177./21582440198988-49

Paloutzian, R. F., and Park, C. L. (Eds). (2013). Handbook of the psychology of religion and spirituality, 2nd ed. Guilford Press, New York, London

Pauen, M., and Haynes, J-. D. (2021). Measuring the mental. *Conscious Cognition* 90, 103106. doi: 10.1016/j.concog.2021.103106.

Romano, A., Giardini, F., Columbus, S., de Kwaadsteniet, E. W., Kisfalusi, D., Triki, Z., et al. (2021). Reputation and socio-ecology in humans. *Philos. Trans. R. Soc. Lond,. B,. Biol. Sci.* 376, 20200295. doi: 10.1098/rstb.2020.0295

Seitz, R. J. (2022). Beliefs - a challenge in neuropsychological disorders. *J. Neuropsychol.* 16, 21–37. doi: 10.1111/jnp.12249

Seitz, R. J., and Angel, H-. F. (2012). Processes of believing - a review and conceptual account. *Rev. Neurosci.* 23, 303–309. doi: 10.1515/revneuro-2012-0034

Seitz, R. J., and Angel, H. F. (2020). Belief formation - a driving force for brain evolution. *Brain and Cognition* 140, 105548. 1-8. doi: 10.1016/j.bandc.2020. 105548

Seitz, R. J., Angel, H. F., and Paloutzian, R. F. (2022). Statements of believing involve attributions. Academia Letters, Article 4624; doi: 10.20935./AL4624

Seitz, R. J., and Angel, H. F. Paloutzian. (2023). Bridging the gap between believing and memory functions. *Eur J Psychol.* 19, 113–124. doi: 10.5964/ejop.7461

Seitz, R. J., Paloutzian, R. F., and Angel, H. F. (2018). From believing to belief: a general theoretical model. *J Cog Neurosci* 30, 1254–1264. doi: 10.1162/jocn_a_01292

Sharot, T., Kanai, R., Marston, D., Korn, C. W., Rees, G., Dolan, R. J., et al. (2012). Selectively altering belief formation in the human brain. Proc Natl Acad Sci USA 109, 17058–17062. doi: 10.107/pnas/1205828109

Sugiura, M., Seitz, R. J., and Angel, H-. F. (2015). Models and neural bases of the believing process. JBBS 5, 12–23. doi: 10.4236/jbbs.2015.51002

Teehan, J. (2006). The evolutionary basis of religious ethics. *Zygon* 41, 747–773. doi: 10.1111/j.1467-9744.2005.00772.x

Teehan, J. (2014). Cognitive science and the limits of theology. In: Trigg R, Barrett J. (eds). The roots of religion. Exploring the cognitive science of religion. Chapter 10, 167–185. Ashgate, UK. doi: 10.4324/9781315553 207-10

Ülkümen, G., Fox, C. R., and Malle, B. F. (2016). Two dimensions of subjective uncertainty: Clues from natural language. *J. Exp. Psychol. Gen.* 145, 1280–1297. doi: 10.1037/xge0000202

Wang, W-. C., Brashier, N. M., Wing, E. A., Marsh, E. J., and Cabeza, R. (2016). On known unknowns: fluency and the neural mechanisms of illusory truth. *J. Cogn. Neurosci.* 28, 739–746. doi: 10.1162/jocn_a_00923

frontiers | Frontiers in Psychology

TYPE Opinion
PUBLISHED 04 November 2022
DOI 10.3389/fpsyg.2022.942590

Check for updates

OPEN ACCESS

EDITED BY
Anna M. Borghi,
Sapienza University of Rome, Italy

REVIEWED BY
Lluis Oviedo,
Pontifical University Antonianum, Italy

*CORRESPONDENCE
Hans-Ferdinand Angel
ferdinand.angel@uni-graz.at

SPECIALTY SECTION
This article was submitted to Theoretical and Philosophical Psychology, a section of the journal Frontiers in Psychology

RECEIVED 12 May 2022
ACCEPTED 11 October 2022
PUBLISHED 04 November 2022

CITATION
Angel H-F (2022) Pathways and crossroads to creditions: Insights from a retrospective view.
Front. Psychol. 13:942590.
doi: 10.3389/fpsyg.2022.942590

COPYRIGHT
© 2022 Angel. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Pathways and crossroads to creditions: Insights from a retrospective view

Hans-Ferdinand Angel*

Karl Franzens University of Graz, Graz, Austria

KEYWORDS

credition, research development, basic research, model of credition, religiosity

Credition is a neologism derived from the Latin word credere (to believe) and designates processes of believing (Angel, 2013a). In many languages (and esp. in German) the term belief is widely associated with religion and religious beliefs. Indeed, the need for a new term became evident during the so-called Regensburg Symposia (1998–2005) (Angel, 2006a) that were aimed at increasing our understanding of the phenomenon of religiosity (see below) and the dynamics of 'religious beliefs'. Given this background, it is important to emphasize that credition is neither a religious nor a theological term. Rather, it was coined as a psychological term in analogy to other psychological terms including cognition, emotion, and volition. No religion is needed in order to understand "credition," but knowledge about credition may help us to better understand religious beliefs. Although the intention of this article is to point to issues which appeared as crossroads and pathways in the emerging history of creditions, it does not present a chronology of events but focusses on theoretical issues.

Precursory hermeneutics as crossroads to credition

Blind spot time-related beliefs

Talking about belief or credition means initially to talk about notions, i.e., the notion of belief and the notion of credition. Both terms highlight related but different phenomena. Belief has been a topic discussed since Antiquity. It might have contributed to our lack of understanding that something like fluid or temporally evolving believing processes might exist because, at least since late Antiquity and the early Middle Ages, the predominant scientific practice has been to talk about 'belief' as a static entity, i.e., as a noun (Angel, 2022a). But on a linguistic level we must address the relation between a noun and its corresponding verb. To proceed from a noun related understanding of beliefs to an action-based understanding of believing processes that can be expressed in terms of a verb (e.g., while believing) requires a paradigm shift. The paradigm shift that underpins going from understanding beliefs to understanding the processes of believing is a precondition

for the concept of credition that must be elaborated within the intersection of different scholarly fields such as linguistics, epistemology, philosophy of mind, cognitive science, neuroscience, sociology, information theory, psychology, and psychology of religion (Angel et al., 2017).

Language related issues

For understanding beliefs and following the path of the paradigm shift, a crucial issue turned out to be the language in which the shift was discussed. Certain terms are central to the credition concept: meaning, mind, perception, evidence, and representation. These terms could be obstacles to understanding believing processes. This shall be exemplified by two lexemes.

(a) Ancient Greek offers two words to express the notion of the verb 'to believe': δοξάζειν [doxázein] and πιστεύειν [pistéuein]. In both cases the relation between noun and verb can be discussed because both verbs have corresponding nouns: δόξα [dóxa] and πίστις [pístis]. The former has been used since the 4th century to express the correct 'orthodox' Christian faith ['ορθός: orthós: correct]. However, the relation between the noun and verb is awkward in Latin. The Latin translation for 'pístis' is 'fides', from which stems the English term 'faith'. The required paradigm shift is impeded because there are no corresponding verbs for the Latin "fides" or the English "faith." To switch from a substantive expression (fides; faith) to a verbal expression, one must change the wordstem (credere, to believe). Another obstacle is that for the English verb "to believe" the noun "belief" exists, whereas the Latin term "credere" lacks a correspondent noun. This lack contributed to the blind spot because until the Renaissance Western philosophy was based on the use of Latin. Also, the English language has two nouns – "belief" and "faith." In contrast, the German language provides only one noun ("Glaube") which covers the semantic broadness of both English nouns. Because there is no verb for the English "faith," often the adjective 'religious' is used to express "having faith."

(b) Specific links between "religion," "faith," and "belief" are apparent. The semantic broadness of 'religion' leads to an impervious terminological mess which deeply infects the research on credition. Research on religions and their role in societies began to flourish in the late 1800's in different ways. After Darwin (1859) an interest in the evolution of religion was fostered (Feierman and Oviedo, 2019); the psychology of religion beginnings included neuropsychological perspectives (James, 1902); and the sociology of religion started to examine the social role of religions (Durkheim, 1912). The 19th-century-debates spread the term religion widely, contributing to its present dominant appeal (Seitz and Angel, 2014). This is apparent in the names of certain scholarly sub-disciplines such as history of religion, psychology of religion, sociology of religion, philosophy of religion, and phenomenology of religion. This predominance of 'religion' causes at least three problems including (1) a marginalization of the term religiosity, (2) the absence of an academic goal to clarify the terms religiosity or religiousness (Angel, 2013b), and (3) the absence of 'religiosity' as theoretical starting point so that many important issues in understanding religious behavior – be it dysfunctional or not – cannot be addressed in a theoretically adequate manner (Seitz et al., 2021). But any theoretically sound understanding of 'religious experience' has to encompass three elements – religion, religiosity, and the individual or collective relation between religion and religiosity (Angel, 2019). Importantly, the nucleus of all later development of the idea of credition is routed in the German language. The Regensburg Symposia helped us to better understand the German term "Religiosität," not "Religion" (Angel, 2006a). "Religiosität" as typical German term cannot be adequately translated into English because there exist three terms – religiosity, religiousness, and spirituality – none of which is fully equivalent to the German term Religiosität. The book-title 'Geschichte der Religiosität im Mittelalter' (Angenendt, 2009) cannot be translated into English in a satisfying manner.

Semantic of (religious) belief(s)

A second – rarely addressed (Sharpe, 1983, p. VIII)"–'problem is the neglect of the linguistic nature of "religious" as an adjective. 'Religious' as an adjective has dual associations: it can be related to two nouns – "religion" and "religiosity." Because associating religious with religion is widely accepted in empirical scientific research, the dual character of the adjective is less apparent. A striking example is the adjective interreligious. Because it is typically associated with different religions, its other function is often lost so that it is seldom invoked to appreciate the different features of religiosity or religiousness as might be possible when considering, for example, open-mindedness vs. fundamentalism. "The common language use seems to put the terms religious, religiosity, and religion in a melting pot from where the words can be taken out in an exchangeable manner" (Seitz et al., 2021, p. 62).

In recent years linguistic philosophy has pointed to the important role of languages in the production of worldviews (Rorty, 1967; Waismann, 1997). The role of languages is also of crucial importance for the interdisciplinary and global credition research project and was prominent in the attempts to conceptualize credition. Thus, it is not mere storytelling when the complicated linguistic issues in the topic of beliefs in general and specifically religious beliefs are highlighted. Three types of issues emerged in relation to the languages used.

(a) It is possible to clarify the relation between nouns and verbs in Indo-European languages, but not in all languages. There are restrictions in generative grammar and

its later developments (Chomsky, 1965, 1986). More advanced ontogenetically based linguistic theories (Tomasello, 2003, 2008) must be integrated into credition research (Seitz et al., 2018), and the role of participles influences the linguistic possibilities. In an Anglo-American but not in a German context, 'believing' can be used in the same manner as is 'learning', prompting discussion of how the cognitive processes of believing and learning are related.

(b) The chaotic religious semantics reflects the ambiguity of its emotional loading. At least in the context of Western thinking it might be adequate to conceive of religion as "an incredibly powerful catalyst for both our best and worst" (Sapolsky, 2017, p. 621). This ambiguity exposes the emotional loading of 'religious' beliefs, and thus the topic of belief in general.

(c) The English 'belief' can be used in a plural form (beliefs), whereas no plural exists for the German "Glaube." This distinction is often explicitly highlighted when epistemic texts are to be translated. For the translation of 'belief' into German, sometimes the term "Meinung" is used instead of "Glaube" (Bieri, 1987, p. 106). But since "Meinung" also conveys the English equivalent "opinion," an identity of belief and opinion is implied which makes it difficult to convey the role of trust in believing. Such a translation follows the Latin speaking trend in philosophy, as the Latin term "credere" does not include the notion of trust, whereas the Greek term "pisteuein" does. This trend, rooted in Latin, tends to end up in the field of epistemology.

(d) Scientific research is often driven by a WEIRD (Western, educated, industrialized, rich, democratic) perspective (Henrich et al., 2010). When developing credition in a globalized context this restricted perspective must be overcome. We need to develop sensitivity to the richness and different mental and emotional roots of non-European languages, as van Leeuwen demonstrates when he compares Fante, Thai, and Mandarin (van Leeuwen et al., 2022).

Milestones of hermeneutic clarification

The research on credition happens in the collaboration of different disciplines with different methodologies and language rules (Wittgenstein, 1953). A considerable amount of preliminary hermeneutic clarifications is needed to comprehend the theoretical groundwork that underpins the credition concept and its neurophysiological base. Learning about credition might appear to be a challenge (Madzarevic, 2022).

(a) "Conceptional questions antecede matters of truth and falsehood. [...] Hence conceptual questions are not amendable to scientific investigation and experimentation. [...] Distinguishing conceptual questions from empirical ones is of first importance" (Bennett and Hacker, 2007, p. 2). Many terms which are embedded in concepts are relevant for an understanding of credition, such as process, function, action, relation (Seitz et al., 2018, p. 1257f.), normal and normality, meaning, value, will and free will, decision and decision-making, and others.

(b) The most adequate synonym for credition seems to be process of believing. Nevertheless, this is correct only in comparison to a static and noun-related understanding. Process is the antonym to state. But in the context of cognitive neuroscience, credition processes must be differentiated from the functions of credition that are attributed to these processes. Here process is the antonym to function.

(c) The term credition can be found in both singular and plural forms in the literature. In its singular use it designates a generic term in analogy to cognition, emotion, volition, and similar terms. When it is used in plural, the intention is to point to the neurophysiological processes that are occurring while someone is believing (Angel, 2022a).

(d) The relation between belief and process of believing can be expressed in mathematical terms: $B = f(b,t)$. This means that belief (B) is a function of believing (b) and the character of a 'belief' depends on what has occurred across the time (t) (Seitz et al., 2018, p. 1257).

Crossroads to an understanding of creditions as brain function

Blind spot neural believing processes

Credition as an idea emerged during the Regensburg Symposia. It was inspired by ongoing debates about the origin of religious beliefs. In cognitive neuroscience two seemingly incompatible and camp-building positions which seemed to be based on two different psychological concepts were held. The limbic marker theory suggests "that the primary substrate for this <religious and mystical; HFA> experience is the limbic system" and "predicts that functional neuroimaging during numinous experiences in individuals who have repeated religious transports would reveal alterations in limbic system activity" (Saver and Rabin, 1999, p. 204). In contrast, a cortical marker theory suggests "that religious experience may be a cognitive process, mediated by a pre-established neural circuit, involving dorsolateral prefrontal, dorsomedial frontal and medial parietal cortex" (Azari et al., 2001, p. 1651). For understanding "Religiosität" cognition and emotion appeared as insufficient categories and believing processes (i.e., creditions) were postulated (Angel, 2006b, p. 71).

The idea of credition was then not more than a postulate, but it allowed us to address believing processes by means of neuroscientific approaches (Seitz, 2017, p. 2022). The paradigm-shift toward understanding beliefs as manifestations of processes faced a similar situation because beliefs had typically been analyzed hermeneutically. A PubMed-review of empirical findings revealed "a lack of empirical effort to understand belief"

(Seitz and Angel, 2012) which is sometimes addressed even as "neglect of belief" (Connors and Halligan, 2015).

Milestones toward credition

Beliefs differ from knowledge because they imply subjective meaning. Thus, one key issue for understanding believing is centered on the role of emotional valuations and subjective meaning-making. A series of publications emphasizes relevant aspects of this (Angel, 2022b).

(a) Like other cognitive processes, the process character of credition includes several different mental operations that are heavily involved in the perception of events or objects in the outer world and in control of behavior (Angel and Seitz, 2016). As shown in Figure 1, this multifunctionality can be specified (Angel, 2017). The so-called enclosure function denotes the self-organizing probabilistic assembly of mental attributes of a given object or event that a person is encountering into a coherent mental construct (Angel and Seitz, 2016). Beliefs can lead to action (converter function) and are stabilized by reinforcement learning (stabilizer function). These supramodal functions are modified by the individuality of agents (modulator function). The "functional anatomy" of the believing process can be described at a neurophysiological level (Seitz, 2017).

(b) Believing can be explained by a dual-component model which combines self-organization process of cognitive and emotional elements with a belief evaluation component. Subjective representations encompass self-cognition that refers to a multi-layered self on a physical, interpersonal, and higher social level. A major role plays the right dorsolateral prefrontal cortex (DLPFC) and the medial frontal cortex (MFC) (Sugiura et al., 2015).

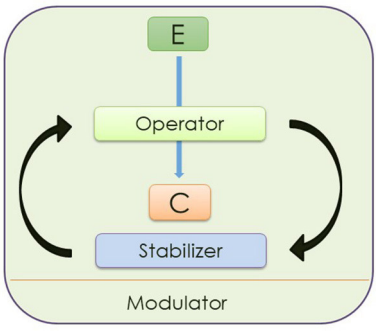


FIGURE 1
The credition model describing the process of believing. The "Enclosure Function" (E) defines the enclosed representation of the perceived stimulus, the "Converter Function"(C) provides the appropriate action in response to the stimulus. By reinforcement learning the putative beliefs are stabilized which is indicated by the "Stabilizer Function." These three supramodal functions are modulated by the internal state of the individual–called "Modulator Function." In the figure the different type of the modulator function is indicated by a thin line. © HF Angel; conference presentation 2012, for the first time published in SFU Research Bulletin, 3/1, 1–20 Angel and Seitz (2016).

(c) To connect the neuroscientific aspect with general anthropological dimensions of believing, the role of emotions in meaning-making was included (Paloutzian and Mukai, 2017). It is suggested that the formation of belief systems and their behavioral consequences can be predicted as result of a probabilistic perception-action-valuation model which represents the mental operations that seem to underly believing processes (Seitz et al., 2016).

(d) Beliefs are "the neuropsychic product of fundamental brain processes that attribute affective meaning to concrete objects and events, enabling individual goal setting, decision making and maneuvering in the environment" which can be categorized as empirical, relational, and conceptual beliefs. "Whilst empirical beliefs about objects and relational beliefs about events develop below the level of awareness and are updated dynamically conceptual beliefs are more complex as being based on narratives and participation in ritual acts" (Seitz and Angel, 2020, p. 1).

(e) This allows us to hypothesize that the 'capacity of believing' is a result of the evolution of the brain. The parietal cortex which accommodates in close vicinity the neural representations of executive, perceptual, and higher order conceptual functions may be a candidate area (Seitz, 2022).

(f) Beliefs are constantly adjusted by the perception of new signals in a Bayesian sense and can be explained as result of believing processes which include learning. They take place on a neurological level but integrate information which have been perceiving from the social environment. Note, the general model results in a mathematically expressed equation:

$$B = S/N \times V + (\alpha \times \delta) \times V_{\delta}. \text{ [Seitz et al 2018, 1259].}$$

(g) From a clinical perspective believing processes can become dysfunctional (Seitz, 2021). This may play a role in psychiatric contexts (Paloutzian et al., 2018) and have an impact on religious beliefs (Seitz et al., 2021).

(h) Maintaining beliefs is interwoven with memory functions in a multifaceted fashion. For instance, linking the typically rapid and adequate reactions of a person to what he or she believes is enabled by working memory. Perceptions are stored in episodic memory as beneficial or aversive events, while the corresponding verbal descriptions of what somebody believes are stored in semantic memory. After recall from memory of what someone believes, personally relevant information can be communicated to other people (Seitz et al., 2022a).

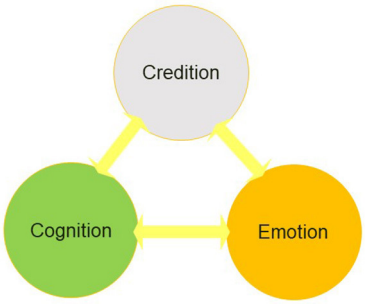


FIGURE 2
The figure shows the interdependence of credition, emotion, and cognition as originally depicted in the basic model of credition. [© HF Angel].

(i) The Credition project follows three research strands: basic (2011), applied (2014), and implementation (2016). The CreditionLab (opened 2018 at the University of Technology in Graz) tests the so called 'model of credition' as specific communication tool intended to make visible the functionality of believing (Angel and Seitz, 2016; Angel, 2017) and seems successfully applied as reference tool for communication-settings (Mitropoulou, 2017; Mitropoulou et al., 2018; Hick et al., 2020; Kranabitl et al., 2021; Lumbreras et al., 2021; Tietz et al., 2022).

Discussion

During the Regensburg Symposia it became necessary to establish credition linguistically as scientific term (Angel, 2006a), although from the beginning creditions were phenomenologically conceived as intertwined with cognition and emotion (Angel, 2016, 2021) (Figure 2: basic model).

Because credition is now a widely used term, the relations between creditions and other emotional and cognitive processes can be addressed (Seitz et al., 2018). As believing is intimately linked with inferential information processing, information that is processed and/or modified in the brain will be labeled with diverse attributions. Typically, these attributions correspond to meta-cognitive self-attributions or third-persons attributions concerning behavior observed in other people (Seitz et al., 2022b). Such (*post-hoc*) attributions are conceptually different from the belief categories that have been defined with respect to the type of information processed (Seitz and Angel, 2020).

To construct the "model of credition," conceptional neurophysiological findings about believing were translated into model-specific terms. Their adequateness and the methodical transformation of the underlying concepts may be discussed. For instance, the production process was inspired by neurophysiological findings about the simultaneous production of cognitive and emotional processes in the prefrontal-medial cortex (Gray et al., 2002; Schaefer and Gray, 2007). Since no term existed to express linguistically this simultaneity "bab" as basic term of the model of credition was coined. It designates: "emotional loaded proposition."

For stable beliefs it might be adequate to talk about religious, political, or economic beliefs. From a processual perspective, that is from a credition perspective, such characteristics do not make sense (Angel 2022a, 615–621).

Author contributions

The author confirms being the sole contributor of this work and has approved it for publication.

Funding

This paper was funded by Dr. Rüdiger Seitz, *via* the Volkswagen Foundation, Siemens Healthineers, and the Betz Foundation.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Angel, H.-F. (2006a). "Religiosität," in: *Religiosität*, eds H.-F. Angel, M. Bröking-Bortfeldt, U. Hemel, M. Jacobs, J. Kunstmann, M. L. Pirner, and M. Rothgangel (Stuttgart: Kohlhammer).

Angel, H.-F. (2006b). "Religiosität als menschliches Potenzial. Ein anthropologisches Modell im neurowissenschaftlichen Horizont," in: *Religiosität*, ed H.-F. Angel (Stuttgart: Kohlhammer).

Angel, H.-F. (2013b). "Religiosity," in: *Encyclopedia of Sciences and Religions*, eds A. L. C. Runehov and L. Oviedo (Dordrecht: Springer).

Angel, H.-F. (2016). No believing without emotion: The overlapping of emotion and cognition in the model of credition. *Stud. Sci. Theol.* 15, 215–222.

Angel, H.-F. (2017). "Credition: From the question of belief to the question of believing," in: *Processes of Believing: The Acquisition, Maintenance, and Change in Creditions*, eds H.-F. Angel, L. Oviedo, R. F. Paloutzian, A. L. C. Runehov, and R. J. Seitz (Cham: Springer). doi: 10.1007/978-3-319-50924-2_2

Angel, H.-F. (2019). "A history of the evolution of religion: from religion to religiosity to the processes of believing," in: *The Evolution of Religion: How Biology, Psychology, and Theology Interact*, eds J. Feierman and L. Oviedo (New York, NY: Routledge/Taylor & Francis). doi: 10.4324/9780429285608-6

Angel, H.-F. (2021). "A process of merging the interior and exterior reality: a short view on the structure of credition," in: *Mind in Nature. Bridging Process Philosophy and Neoplatonism*, ed. M.-T. Teixeira, M. Reker, and A. Berve (Cambridge: Scholars Publishing).

Angel, H.-F. (2022a). *Credition. Fluides Glauben. Kultur- und Wissenschaftsgeschichte von einem blinden Fleck und seinem Ende*. (Baden-Baden: Deutscher Wissenschaftsverlag).

Angel, H.-F. (2022b). *Credition*. Available online at: https://credition.uni-graz.at/de/archive/ (accessed July 26, 2022).

Angel, H.-F., Oviedo, L., Paloutzian, R. F., Runehov, A. L. C., and Seitz, R. J. (2017). *Processes of Believing: The Acquisition, Maintenance, and Change in Creditions*. Cham: Springer. doi: 10.1007/978-3-319-50924-2

Angel, H.-F., and Seitz, R. J. (2016). Process of Believing as Fundamental Brain Function: The Concept of Credition. *SFU Res. Bullet.* 3, 1–20. Available online at: https://static.uni-graz.at/fileadmin/Credition/Creditions_2016/CredAngelSeitz2016_1_.pdf

Angel, H. F. (2013a). "Credition, the process of belief," in: *Encyclopedia of Sciences and Religions*, eds A. L. C. Runehov and L. Oviedo (Dordrecht: Springer). doi: 10.1007/978-1-4020-8265-8_1565

Angenendt, A. (2009). *Geschichte der Religiosität im Mittelalter*. Darmstadt: Wiss. Buchgesellschaft.

Azari, N. P., Nickel, J., Wunderlich, G., Niedeggen, M., Hefter, H., Tellmann, L., et al. (2001). Neural correlates of religious experience. *Euro. J. Neurosci.* 13, 1649–1652. doi: 10.1046/j.0953-816x.2001.01527.x

Bennett, M. R., and Hacker, P. M. S. (2007). *Philosophical Foundations of Neuroscience. 7th edition*. Malden: Blackwell.

Bieri, P. (1987). *Analytische Philosophie der Erkenntnis*. Frankfurt: Athenäum.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press. doi: 10.21236/AD0616323

Chomsky, N. (1986). *Knowledge of Language. Its Nature, Origin and Use*. New York, NY: Praeger.

Connors, M. H., and Halligan, P. W. (2015). A cognitive account of belief: A tentative roadmap. *Front. Psychol.* 5, 1588. doi: 10.3389/fpsyg.2014.01588

Darwin, C. (1859). *On the Origin of Species by Means of Natural Selection*. London: John Murray.

Durkheim, É. (1912). *Les formes élémentaires de la vie religieuse. Le système totémique en Australie*. Paris: Félix Alcan.

Feierman, J., and Oviedo, L. (2019). *The Evolution of Religion: How Biology, Psychology, and Theology Interact*. New York, NY: Routledge/Taylor & Francis. doi: 10.4324/9780429285608

Gray, J. R., Braver, T. S., and Raichle, M. E. (2002). Integration of emotion and cognition in the lateral prefrontal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 99, 4115–4120. doi: 10.1073/pnas.062381899

Henrich, J., Heine, S., and Norenzayan, A. (2010). The weirdest people in the world? *Behav. Brain Sci.* 33, 61–83. doi: 10.1017/S0140525X0999152X

Hick, H., Angel, H.-F., Kranabitl, P., and Wagner-Skacel, J. (2020). "Decision making and the influence of the human factor," in *Systems Engineering for Powertrain Development*, eds H. Hick, K. Küpper, H. Sorger (Cham: Springer). doi: 10.1007/978-3-319-68847-3_14-1

James, W. (1902). *The Varieties of Religious Experience: A Study in Human Nature*. London: Longmans, Green and Co. doi: 10.1037/10004-000

Kranabitl, P., Faustmann, C., and Hick, H. (2021). Decision making for sustainable technical applications with the SMH approach. *Sustainability* 13:8702. doi: 10.3390/su13168702

Lumbreras, S., Oviedo, L., and Angel, H.-F. (2021). The missing piece in sustainability indices: accounting for the human factor. *Sustainability* 13:21. doi: 10.3390/su132111796

Madzarevic, D. (2022). Learning about Credition: Exploring the barriers between basic and applied research. *Front. Behav. Neurosci*. doi: 10.3389/fpsyg.2022.943376

Mitropoulou, V. (2017). "Understanding young people's worldview: a practical example of how to work with the model of credition," in: *Processes of Believing: The Acquisition, Maintenance, and Change in Creditions*, ed H.-F. Angel (Cham: Springer). doi: 10.1007/978-3-319-50924-2_25

Mitropoulou, V., Gkirlou, D., and Meke, M. (2018). Building worldviews with credition teaching model in high school curricula of religious education: a controlled experimental study. *Sociol. Stud.* 8:7. doi: 10.17265/2159-5526/2018.07.003

Paloutzian, R. F., Angel, H.-F., and Seitz, R. J. (2018). The process of believing and psychiatric symptoms. *Relig. Brain Behav*. 10:184. doi: 10.1080/2153599X.2018.1532456

Paloutzian, R. F., and Mukai, K. J. (2017). "Believing, remembering, and imagining: the roots and fruits of meanings made and remade," in: *Processes of Believing: The Acquisition, Maintenance, and Change in Creditions*, ed H.-F. Angel (Cham: Springer). doi: 10.1007/978-3-319-50924-2_3

Rorty, R. (1967). *The Linguistic Turn: Recent Essays in Philosophical Method*. Chicago, IL: University of Chicago Press.

Sapolsky, R. M. (2017). *Behave: The Biology of Humans at Our Best and Worst*. London: Penguin Press.

Saver, J. L., and Rabin, J. (1999). "The neural substrates of religious experience," in: *The Neuropsychiatry of Limbic and Subcortical Disorders*, eds S. Salloway, P. Malloy, J. L. Cummings (Washington and London: Amer Psychiatric Pub Inc.).

Schaefer, A., and Gray, J. R. (2007). A role for the human amygdala in higher cognition. *Rev. Neurosci.* 18, 355–364. doi: 10.1515/revneuro.2007.18.5.355

Seitz, R. J. (2017). "Beliefs and believing as possible targets for neuroscientific research," in: *Processes of Believing: The Acquisition, Maintenance, and Change in Creditions*, ed H.-F. Angel (Cham: Springer). doi: 10.1007/978-3-319-50924-2_5

Seitz, R. J. (2021). Beliefs: A challenge in neuropsychological disorders. *J. Neuropsychol.* 16, 21–37. doi: 10.1111/jnp.12249

Seitz, R. J. (2022). Believing and Beliefs—Neurophysiological Underpinnings. *Front. Behav. Neurosci*. 16:880504. doi: 10.3389/fnbeh.2022.880504

Seitz, R. J., and Angel, H.-F. (2012). Processes of believing - a review and conceptual account. *Rev. Neurosci.* 23, 303–309. doi: 10.1515/revneuro-2012-0034

Seitz, R. J., and Angel, H.-F. (2014). Psychology of religion and spirituality. Meaning-making and processes of believing. *Relig. Brain Behav*. 15, 12–23. doi: 10.1080/2153599X.2014.891249

Seitz, R. J., and Angel, H.-F. (2020). Belief formation – A driving force for brain evolution. *Brain Cognit.* 140:105548. doi: 10.1016/j.bandc.2020.105548

Seitz, R. J., Angel, H.-F., and Paloutzian, R. F. (2021). The Process of believing, mental abnormalities, and other matters of the mind: where do they come from? What are they good for? *J. Cognit. Sci. Relig.* 7, 54–72. doi: 10.1558/jcsr.19559

Seitz, R. J., Angel, H.-F., and Paloutzian, R. F. (2022a). Bridging the gap between believing and memory functions. *Europe's J. Psychol*.

Seitz, R. J., Angel, H.-F., and Paloutzian, R. F. (2022b). Statements of believing involve attribution. *Acad. Lett.* 2022:4624. doi: 10.20935/AL4624

Seitz, R. J., Paloutzian, R. F., and Angel, H.-F. (2016). Processes of believing: Where do they come from? What are they good for? *F1000 research* 5:e9773. doi: 10.12688/f1000research.9773.1

Seitz, R. J., Paloutzian, R. F., and Angel, H.-F. (2018). From believing to belief: A general theoretical model. *J. Cognit. Neurosci.* 30, 1254–1264. doi: 10.1162/jocn_a_01292

Sharpe, E. J. (1983). *Understanding Religion*. London: Duckworth.

Sugiura, M., Seitz, R. J., and Angel, H.-F. (2015). Models and neural bases of the believing process. *J. Behav. Brain Sci.* 5, 12–23. doi: 10.4236/jbbs.2015.51002

Tietz, S., Wagner-Skacel, J., Angel, H.-F., Ratzenhofer, M., Fellendorf, F. T., Fleischmann, E., et al. (2022). Believing processes during the COVID-19 pandemic in individuals with bipolar disorder: An exploratory study. *World J. Psychiatry* 12, 929–943. doi: 10.5498/wjp.v12.i7.929

Tomasello, M. (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press.

Tomasello, M. (2008). *Origins of Human Communication*. Cambridge, MA: MIT Press. doi: 10.7551/mitpress/7551.001.0001

van Leeuwen, N., Weisman, K., and Luhrmann, T. M. (2022). To believe is not to think: a cross-cultural finding. *Open Mind* 5, 91–99. doi: 10.1162/opmi_a_00044

Waismann, F. (1997). *The Principles of Linguistic Philosophy*. New York, NY: Palgrave Macmillan. doi: 10.1007/978-1-349-25335-7

Wittgenstein, L. (1953). *Philosophical Investigations*. New York, NY: Macmillan Publishing Company.

frontiers | Frontiers in Psychology

TYPE Opinion
PUBLISHED 29 July 2022
DOI 10.3389/fpsyg.2022.981925

Check for updates

OPEN ACCESS

EDITED BY
Rüdiger J. Seitz,
Heinrich Heine University of Düsseldorf, Germany

REVIEWED BY
Peter Holzer,
Medical University of Graz, Austria

*CORRESPONDENCE
Mahault Albarracin
mahault.albarracin@gmail.com
Riddhi J. Pitliya
c-riddhi.jain@verses.io

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Theoretical and Philosophical Psychology, a section of the journal Frontiers in Psychology

RECEIVED 30 June 2022
ACCEPTED 13 July 2022
PUBLISHED 29 July 2022

CITATION
Albarracin M and Pitliya RJ (2022) The nature of beliefs and believing. *Front. Psychol.* 13:981925. doi: 10.3389/fpsyg.2022.981925

COPYRIGHT
© 2022 Albarracin and Pitliya. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

The nature of beliefs and believing

Mahault Albarracin[1,2*†] and Riddhi J. Pitliya[1,3*†]

[1]VERSES Research Lab, Los Angeles, CA, United States, [2]Département d'informatique, Université du Québec à Montréal, Montréal (Québec), QC, Canada, [3]Department of Experimental Psychology, University of Oxford, Oxford, United Kingdom

KEYWORDS

active inference, beliefs (non-religious), epistemic (belief) changes, socio-political, psychopathology

The nature of beliefs

Conceptualizations of beliefs differ according to the school of thought considered; here, we take the view from cognitive science.

In cognitive science, beliefs are propositional attitudes, where the world is depicted as being in some state or another (Schwitzgebel, 2021). Beliefs have two main properties: some representational content and assumed veracity (Stephens and Graham, 2004). Beliefs entail specific representational content, which portrays causes of sensations (agency, events, and objects) as being a specific way (Rimell, 2021). So understood, they are undoubtedly a central part of cognition, dictating our perceptions, behavior, and executive functions. Beliefs do not need to be conscious or linguistically articulated, and indeed, the majority of beliefs can be construed as subpersonal; i.e., remain unconscious (Majeed, 2022). Rational agents generally view beliefs as having a truth value, and update their beliefs in light of new evidence. The term "belief" is also used to denote a more deflationary sense, where what is at stake is merely a probability density over some support; where we call a belief a probabilistic assessment of how plausible some state of affairs is (Smets, 2005). On this probabilistic reading, beliefs acquire the attribute of uncertainty—or its complement precision.

Beliefs provide the foundation that allows agents to understand—or at least make sense of—the world and act within it: they provide agents with a consistent and coherent representation of their world, which they can then use to make inferences about the causal structure of the world and their place within it (Churchland and Churchland, 2013). This scaffolding of beliefs helps [human] agents appraise the environment, explain new observations, construct shared perspectives on the world, and engage in goal-directed behavior.

Beliefs also help us experience the world temporally, as they can represent the state of the world in the past and allow us to anticipate its state in the future; this is especially important when holding beliefs about the consequences of action—a prerequisite for planning and a sense of agency (Shipp et al., 2009).

Active inference

Active inference is a formal description of self-organization derived from the variational free energy principle, and provides a mechanistic account of belief-guided

action (Friston, 2010; Friston et al., 2017). In particular, it is used to model and simulate how beliefs about the states of the world are formed and updated. It proposes to naturalize belief in the first sense by formalizing them as beliefs in the second sense (Friston et al., 2015). In this context, the semantic content of a belief is equated with the support of its probability density (i.e., the external or worldly states over which it is defined). Figure 1 illustrates the dynamic relation between (internal, external, and blanket) states.

Belief formation

Active inference rests on a particular partition of a system (i.e., into particles), where we distinguish between the states internal and external to a system, as well as the states that constitute the boundary of the system via which it interacts with its environment (Friston, 2010; Friston et al., 2017). The external states are assumed to be hidden from direct observation behind the blanket states. In other words, sensory states allow the agent to access their environment vicariously, by sampling it through action. Internal states encode beliefs about the external world, in the form of sufficient statistics (e.g., an expected value or expectation and the confidence of precision of that belief). The agent thus forms beliefs about the hidden states of the external world that have caused their observations (sensory data). Through this accumulation of observations, agents continuously update their beliefs about external states of affairs, and about the most likely future states—that clearly depend upon action.

Belief updating

These beliefs are not static. Depending on how well their beliefs enable them to predict the world, agents can update their (Bayesian) beliefs about of the world. This can be read as a Bayesian mechanics in which, agents acuminate and assimilate new evidence, and re-calibrate what they believe to be the cause of their sensory experiences (Ramstead et al., 2022). This process is generally read as minimizing surprise (technically, the negative logarithm of model evidence of observations). This is mathematically the same as maximizing the evidence for a model of the world; sometimes called a world model or generative model—that generates predictions of (observable) sensory states from their (hidden) external causes.

Active states allow the agent to interact with her world, by directly affecting the process generating her sensory data (i.e., the environment), and accordingly update her beliefs about the external world (Of course, new observations may be formed by causes other than the agent herself) (Friston et al., 2016). This enactive aspect of inference now affords the opportunity to plan responses and choose those actions that will minimize surprise and maximize model evidence; sometimes referred to as self-evidencing.

In active inference, policy selection determines the agent's actions. A belief about a current course of action is called

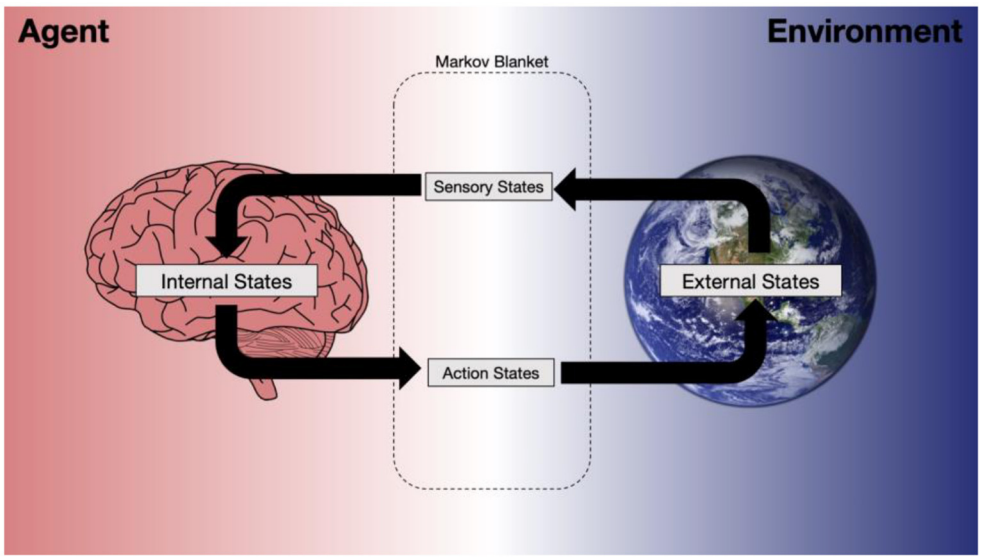


FIGURE 1
An illustration of how a model of the external world is formed, resulting in beliefs about the true states of the world and belief updating, encoded by internal states.

a "policy." In this setting, the agent also forms beliefs about what she is doing, and updates her beliefs about the current course of action (i.e., its policy) based on her other beliefs about the current state of world and the goals that the agent is trying to achieve. The different policies available to an agent are then variations of beliefs about expected future observations, contingent on different courses of action. A policy is selected based on preferences for specific sensory outcomes, by choosing the one that minimizes the divergence between expected and preferred outcomes. This divergence between expected and preferred outcomes is quantified using "expected free energy." The optimal policy is the one that provides the most evidence for the generative model (or equivalently, that is expected to generate the least expected surprise or uncertainty). Active inference thereby purports to explain the dual aspect of choice behavior; namely preference-seeking and information-seeking behavior, respectively. Both aspects are in the service of avoiding surprises (e.g., non-preferred, outcomes). Technically, expected free energy can be decomposed into expected information gain and expected loss; in the sense of optional Bayesian design and Bayesian decision theory, respectively.

Perception is realized by a process called state estimation (Oliver et al., 2021). Beliefs about the current state of the world are updated, based on priors that encode what the agents believes to be the base rates of occurrence of various states of affairs, the likelihood of current observations being caused by certain states, and the probability that states will transition to other states. All knowledge about past states is implicitly packed into beliefs about the current state: belief updating of this sort satisfies what is known as the Markov property and is the basis of all evidence accumulation and inference. The structures that underlie this kind of belief updating are themselves optimized to maximize model evidence leading to an understanding of learning in terms of optimizing the parameters of a generative model (e.g., learning the weights of a neural network or associative plasticity in neuronal networks).

Beliefs across scales

The basic mechanics of active inference can be deployed across nested spatial and temporal scales, describing a nested set of agents composed of other agents—e.g., organs composed of cells and communities composed of individuals—thanks to its scale invariant construction (Kirchhoff et al., 2018). Each scale relies on a formulation of self, which is bounded in a way that individuates self from non-self.

These boundaries just are the Markov blankets above: a Markov blanket comprises of the sensory and action states that separates the internal states from external states of a system (see Figure 1). Markov blankets are drawn at the boundary of a system and define the interface where it receives sensory stimulation. Thus, depending on the perspective, an agent can have its own Markov blanket, or can be a part of a higher-order blanketed system, as a component of that system (Kirchhoff et al., 2018).

Shared beliefs

A Markov blanket can also be drawn over a group of agents. These agents gain an advantage, namely, they can pool their resources, gathering evidence as a group without having to expend the effort necessary to acquire in isolation. The individual agents forming a higher-order agent thereby improve their own predictive abilities and minimize expected free-energy as a group, aligning their beliefs as they share their beliefs about the (co-constructed) world. In the ideal case, agents generate evidence for a shared model of the world. The more similar each agent's model is, the more likely they are to minimize surprise, predict each other and resolve uncertainty, i.e., jointly self-evidencing. Thus, sampling from members of a group that share a similar model itself provides significant informational and evolutionary benefits and allows groups of agents to optimize their beliefs over longer timescales than would otherwise be possible; c.f., cultural eco-niche construction (Bouizegarene et al., 2020).

Applications

We can use the definition of beliefs as propositions of the true—but unknowable—states of the world, and the formalization of belief formation and updating as proposed by active inference, to understand several socio-political and psychopathological phenomena.

Socio-political phenomena

When two agents interact, the value of their beliefs is weighed against the beliefs and observations to which the individual agents already have access. It is advantageous for individuals to be part of a group in terms of large-scale coordination, available resources, and computational power. But it is costly to change any single agent's deeply held beliefs. Therefore, individual agents will give more weight to sensory evidence from agents who are believed to share the same generative model or world narrative. The more agents in a group consider some data points as valuable, and as constituting evidence for their beliefs, the more likely it is that the agents in that group will conform to the beliefs of the group. This is a form of epistemic confirmation bias. In extreme cases, agents can then fall into "echo chambers," in which their beliefs no longer align with any other interpretation of the world, outside that of their own group. We can see this phenomenon at play with

misinformation on social media, spreading extremely fast in echo-chambers, and leading to radicalization (Albarracin et al., 2022).

When a group exists long enough and shares a core set of beliefs, it tends to leave traces in the environment that reflect those beliefs. Consider agents walking in a park. As they walk on the grass, they tend to leave footprints, where the grass will be less likely to grow. Agents noticing these traces may be likely to follow a similar path, reinforcing the path in the grass. This recursive process engraves shared beliefs in the structure of the environment, serving as an anchor around which narratives and semantics will be carried through generations of agents in similar groups. As agents continue to produce shared beliefs, they become embedded in a culture and a cultural materiality, which allows them to predict—and navigate—each other and their environment more easily (Ramstead et al., 2020).

Psychopathology

Active inference is often used in computational psychiatry, to characterize psychiatric conditions as a consequence of atypical beliefs. Strange and untrue beliefs may form based on an individual's experience of the world and the structure of their beliefs, giving rise to psychiatric conditions.

Consider schizophrenia, which has been characterized by a failure of sensory attenuation, with secondary consequences for the dysfunctional acquisition of beliefs about the world and interpersonal interactions (Brown et al., 2013). In the context of sensory exchanges with the world, a failure of sensory attenuation means that the weight of the sensory data (observations) is too high, relative to the prior beliefs about the causes of sensations (Friston et al., 2014). This may explain the inability of schizophrenic patients to infer regular contingencies between stimuli in the world. However, since the prior beliefs are themselves compromised, surprising new data paradoxically improves the detection of new stimuli and contingencies. At the same time, this may also result in such individuals misinterpreting ambiguous information and inferring contingencies that do not exist, resulting in inaccurate beliefs (or delusions). For example, individuals with schizophrenia tend to fail to contextualize the consequences of their actions and thus possess false beliefs about their agency (Jeannerod, 2009). Such individuals may also "jump to conclusions"—i.e., they may require less evidence to form a strong belief (FitzGerald et al., 2015).

If an individual's beliefs cannot be flexibly updated, they may consistently act in a way to align their observations with their beliefs, which would result in dysfunctional behavior. Such individuals would not be able to update their beliefs, even when confronted with new observations that challenge their current beliefs. This is a common symptom observed in individuals with depression, where they tend to not act in their environment even when they do have control, giving rise, over time, to a sense of learned helplessness, and to low motivation and inaction (Grahek et al., 2019).

Summary

Beliefs are propositions about the true states of the world. Active inference—a process theory based on the free energy principle—describes how an agent forms and updates beliefs. The active inference framework posits that the agent (i) observes the world, (ii) infers the causes of the observations, and (iii) forms beliefs about the external states of the world. The agents then act in the world, prompting new observations, and thereby update their beliefs. These beliefs underwrite how an agent approaches the world, and how they will navigate through it, given possible paths into the future. Modeling how agents update their beliefs is thus central to understanding both micro- and macro-phenomena, such as deviations in beliefs at an individual level, resulting in dysfunctional behavior and the development psychiatric conditions. And, at the group level, allowing us to better understand socio-political dynamics in multi-agent scenarios.

Our commentary is an addition to Credition—An Interdisciplinary Approach to the Nature of Beliefs and Believing. In this edition, Seitz, Angel, Paloutzian and Taves have put together a series of perspectives on believing which furthers our understanding of the ways beliefs play a part in cognition, technology and science, ranging from understanding how beliefs are shaped, to their role in artificial intelligence, to the role of culture in beliefs, and more.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Funding

This article was funded by RS, via the Volkswagen Foundation, Siemens Healthineers, and the Betz Foundation and the VERSES computational phenomenology lab.

Acknowledgments

The authors thank Maxwell J. D. Ramstead and Karl J. Friston for useful feedback.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Albarracin, M., Demekas, D., Ramstead, M. J., and Heins, C. (2022). Epistemic communities under active inference. *Entropy* 24:476. doi: 10.3390/e24040476

Bouizegarene, N., Ramstead, M. J., Constant, A., Friston, K. J., and Kirmayer, L. J. (2020). Narrative as active inference: an integrative account of the functions of narratives. *Preprint* 10. doi: 10.31234/osf.io/47ub6

Brown, H., Adams, R. A., Parees, I., Edwards, M., and Friston, K. (2013). Active inference, sensory attenuation and illusions. *Cognitive Proc.* 14, 411–427. doi: 10.1007/s10339-013-0571-3

Churchland, P. S., and Churchland, P. M. (2013). *What Are Beliefs? The Neural Basis of Human Belief Systems*. Hove and New York: Psychology Press, 1–18.

FitzGerald, T. H., Schwartenbeck, P., Moutoussis, M., Dolan, R. J., and Friston, K. (2015). Active inference, evidence accumulation, and the urn task. *Neural Computation* 27, 306–328. doi: 10.1162/NECO_a_00699

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* 11, 127–138. doi: 10.1038/nrn2787

Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., Pezzulo, G., et al. (2016). Active inference and learning. *Neurosci. Biobehav. Rev.* 68, 862–879. doi: 10.1016/j.neubiorev.2016.06.022

Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., and Pezzulo, G. (2017). Active inference: a process theory. *Neural Comput.* 29, 1–49. doi: 10.1162/NECO_a_00912

Friston, K., Rigoli, F., Ognibene, D., Mathys, C., Fitzgerald, T., and Pezzulo, G. (2015). Active inference and epistemic value. *Cognitive Neurosci.* 6, 187–214. doi: 10.1080/17588928.2015.1020053

Friston, K. J., Stephan, K. E., Montague, R., and Dolan, R. J. (2014). Computational psychiatry: the brain as a Phantastic organ. *Lancet Psychiatry* 1, 148–158. doi: 10.1016/S2215-0366(14)70275-5

Grahek, I., Shenhav, A., Musslick, S., Krebs, R. M., and Koster, E. H. (2019). Motivation and cognitive control in depression. *Neurosci. Biobehav. Rev.* 102, 371–381. doi: 10.1016/j.neubiorev.2019.04.011

Jeannerod, M. (2009). The sense of agency and its disturbances in schizophrenia: a reappraisal. *Exp. Brain Res.* 192, 527–532. doi: 10.1007/s00221-008-1533-3

Kirchhoff, M., Parr, T., Palacios, E., Friston, K., and Kiverstein, J. (2018). The markov blankets of life: autonomy, active inference and the free energy principle. *J. Royal Soc. Interface* 15:20170792. doi: 10.1098/rsif.2017.0792

Majeed, R. (2022). The relationship between conscious and unconscious intentionality. *Philosophy* 97, 169–185. doi: 10.1017/S0031819121000383

Oliver, G., Lanillos, P., and Cheng, G. (2021). "An empirical study of active inference on a humanoid robot," in *IEEE Transactions on Cognitive and Developmental Systems*. 14, 462–471. doi: 10.1109/TCDS.2021.3049907

Ramstead, M. J., Friston, K. J., and Hip'olito, I. (2020). Is the free-energy principle a formal theory of semantics? From variational density dynamics to neural and phenotypic representations. *Entropy* 22:889. doi: 10.3390/e22080889

Ramstead, M. J., Sakthivadivel, D. A., Heins, C., Koudahl, M., Millidge, B., Da Costa, L., et al. (2022). *On Bayesian Mechanics: A Physics of and by Beliefs*. arXiv preprint *arXiv:*2205.11543.

Rimell, N. (2021). *Representational Content and the Objects of Thought*. Palgrave Macmillan Singapore: Springer. doi: 10.1007/978-981-16-3517-5

Schwitzgebel, E. (2021). "Belief," in *The Stanford Encyclopedia of Philosophy, Winter 2021 Edition, Metaphysics Research Lab*, ed E. N. Zalta (Stanford University).

Shipp, A. J., Edwards, J. R., and Lambert, L. S. (2009). Conceptualization and measurement of temporal focus: the subjective experience of the past, present, and future. *Organizat. Behav. Human Decision Proc.* 110, 1–22. doi: 10.1016/j.obhdp.2009.05.001

Smets, P. (2005). Belief functions on real numbers. *Int. J. Approximate Reasoning* 40, 181–223. doi: 10.1016/j.ijar.2005.04.001

Stephens, G. L., and Graham, G. (2004). Reconceiving delusion. *Int. Rev. Psychiatr.* 16, 236–241. doi: 10.1080/09540260400003982

frontiers | Frontiers in Psychology

TYPE Opinion
PUBLISHED 28 July 2022
DOI 10.3389/fpsyg.2022.946953

Check for updates

OPEN ACCESS

EDITED BY
Rüdiger J. Seitz,
Heinrich Heine University
of Düsseldorf, Germany

REVIEWED BY
Hans Van Eyghen,
VU Amsterdam, Netherlands

*CORRESPONDENCE
Lluis Oviedo
loviedo@antonianum.eu

SPECIALTY SECTION
This article was submitted to
Theoretical and Philosophical
Psychology,
a section of the journal
Frontiers in Psychology

RECEIVED 18 May 2022
ACCEPTED 29 June 2022
PUBLISHED 28 July 2022

CITATION
Oviedo L (2022) Why does the
multidisciplinary study of beliefs
and believing matter so much?
Front. Psychol. 13:946953.
doi: 10.3389/fpsyg.2022.946953

COPYRIGHT
© 2022 Oviedo. This is an open-access
article distributed under the terms of
the Creative Commons Attribution
License (CC BY). The use, distribution
or reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Why does the multidisciplinary study of beliefs and believing matter so much?

Lluis Oviedo*

Theology Department, Antonianum University, Rome, Italy

KEYWORDS

beliefs, social systems, values, meaning, politics, economy

Introduction

Considerable progress has been made in recent years in the study of beliefs and the believing process. Several programs are trying to better understand this cognitive function and its undeniable contribution to human development and success. The time has now come to better place the dynamics of believing in connection with other cognitive functions and with social systems. Indeed, we are becoming more aware of the important role that beliefs play as a central dimension in human cognition and behavior, about the function of shared beliefs in the stability of social systems and in human interaction and communication. More research is needed to better describe how beliefs and believing contribute to humans dealing with their own environment and other people; to keep working social systems, like the economy, politics, science, education, the judiciary, and obviously religion; and how such sets of beliefs are connected with those social structures. Believing can be observed as a clear case in which the psychological dimension appears as entrenched with the social, rendering those social systems viable; indeed those beliefs appear paramount for the formation of such social systems.

The present short reflection neglects the issue of the role believing plays in general cognition, an issue that has been intensely researched in cognitive sciences and epistemology. In what follows, the focus will lay on the social dimensions linked to belief and believing.

To clarify the concept of belief, some standard dictionaries provide clear definitions; for instance, the *Merriam-Webster* offers the following: "a state or habit of mind in which trust or confidence is placed in some person or thing"; while the *Oxford English Dictionary* offers this one: "a strong feeling that something/somebody exists or is true; confidence that something/somebody is good or right." As can be appreciated, different dimensions converge, which include cognition, which points to the true value, confidence or trust, and finally even goodness.

How important are beliefs for society at large and its sustainability?

A recent article about developments in fundamental physics, published in the weekly news magazine *The Economist*, introduced the topic with the words "By abandoning

some long-held beliefs, physicists are clearing a path to the future" (28 August 2021, p. 63). More recently, the same periodical was published in a section dedicated to economic analysis an article with the title "War and wokery," about how the recent international conditions are pressing for a greater ethical engagement in the economy. The article used four times the words "belief" or "believes," and one "faith." The author states, for instance: "Mr. Sonnenfeld [. . .] has become the high priest of a belief system in Western business. . ." (*The Economist*, 2 April 2022, p. 63). A different case emerges in politics. For example, the recent book of Cass Sunstein, *This is Not Normal: The Politics of Everyday Expectations* (Sunstein, 2021), insists on how much democracy depends on the beliefs of people, beliefs which are quite unstable and changeable, even if in other cases they become more resistant and work as group markers, like the religious ones.

The quoted cases support Agustin Fuentes, who in his recent book *Why We Believe?* (Fuentes, 2019), claimed that science and economics are systems of beliefs too and that this condition invites us to consider both critical spheres of human activity in a different light, or within a specific framework, beyond the certainties and strengths that science and the economy have always claimed. Everything gets a new light and deserves a different treatment when it is assumed that the cognitive model these realities follow falls more on the side of believing than the one of factual and tested knowledge. This step is quite remarkable if we keep in mind that until quite recently, talk or titles about beliefs and believing usually related to religious or supernatural views and deeply held convictions (Shermer, 2011; Shagan, 2021). Indeed, what is surprising in Fuentes's book is that it is not a book about religion, not only.

We are now becoming aware of the extent of belief and how the believing process informs most aspects of human cognition. Indeed, much progress has been achieved when we consider that believing is not only concerned with religious faith and practice, or with broadly held values and meaning. Rather, believing is concerned with a cognitive dimension that is involved in many aspects of human life and social systems. Such a discovery has been brought about by a much more accurate analysis and understanding of the process of believing and a general acknowledgment of the impact of biases in academic life. Such awareness could render beliefs or believing a less reliable cognitive activity, one all too often troubled by deception and delusions. Since Aristotle, believing has been contrasted with knowing, based on strict epistemic methods; a rather second-class cognition, reserved for other areas where the ideal model was harder to attain (Miller, 2013).

Rescuing the meaning and value of believing has not been easy. The previously quoted book of Agustin Fuentes has given us important insight and nourished a new interest in this field. Ongoing projects, like *Creditions*, based in Graz University aimed at researching the belief process, are helping to better clarify that complex process (Castillo et al., 2015; Connors and Halligan, 2015; Angel et al., 2017). What we need now is to better clarify the fields in which beliefs and believing play an essential role, not only a provisional one that could be replaced in short order by more reliable cognitive means.

Before going into the proposed analysis, a thesis can be proposed: beliefs are required as conditions for the formation of every social system, not just religion. The thesis can appear too bold for many, but for others, this is just a truism: without shared beliefs, we cannot conceive how systems like the economy, politics, and the judiciary could work. Some examples will suffice.

Revisiting social systems as believing systems

Science is the first case to consider. Scientists need to hold general beliefs about the world we inhabit, its knowability, and the ability of our theories and models to represent it. Then, when scientists formulate their models based on the available data and analysis, they need to believe that those assumed will work better than alternative ones, something which cannot be taken for granted. This often opens new challenges with data and analysis that could disprove earlier models that most colleagues believed. Pluralism of methods–even in statistical analysis–requires that a researcher puts their faith in one procedure rather than another since choices are unavoidable and so do biases and assumptions. In that sense, a fallibilist model of science, as is the one inspired by Popper, cannot avoid relying on beliefs, more than on certainties. However, it is disputed to what extent scientists just "believe" or rather "know"; obviously in many cases, they know beyond doubt, while in many others their certainty levels come close to believing in the way it has been previously defined.

In the economy, things are more complex, since there are many factors involved in that human activity, and social interaction renders it less predictable. The many crises we have lived through have been not just economic or financial crises, but crises of economic models. Pluralism is present and subtle in economic theory and analysis. In such a panorama, economics as an academic activity depends to a considerable extent on shared basic beliefs and values. The issue becomes still more acute when we deal with real economic subjects: the beliefs and values that they held to determine the course of economic activity; their expectations affect decisions and behaviors. Economic functioning requires trust in other people and institutions, and this is basically a form of belief.

The economic-inspired awareness of the importance of beliefs finds a special application in a related field that now assumes an autonomous status: sustainability studies. In this case, beliefs are clearly involved in any attempt to design sustainable systems applying the standard 3 ESG dimensions:

environmental, social, and governance. We can speak about a human factor deeply ingrained in programs aimed at ensuring a better future for all, or just at the endurance of social bodies. Held beliefs and values are indeed informing the behavior of producers, politicians, and consumers, and those general views will determine whether it is worthwhile to undertake some sacrifices or to pay more attention to measures targeted at saving energy and other resources.

Moving to a different area, beliefs become central in psychotherapy. Indeed, it is broadly assumed that psychological distress and suffering are often linked to wrong beliefs and that some beliefs help to cope with harsh crises, while others usually worsen personal conditions, life quality, and relationships. The point is still more evident when dealing with vaccination campaigns: believing in its efficacy contributed to preventing attitudes of resistance. Moreover, believing the goodness of treatment clearly helps its efficacy. Once more, the human factor needs to be considered in therapeutic processes, besides the usual technical issues and their effects.

Other social systems can be reviewed under those critical lenses revealing them as sets of shared beliefs. This principle applies, for instance, to the judiciary, to the political system that undergoes democracy, to education in all its stages, to the system of media and information, and even to the system of broad social interaction. We need to keep some levels of trust or belief about the reliability of those we meet and those with whom we have exchanged. Trust appears–even in the usual definitions–as a very close concept to "belief," but clearly applied to persons: believing in somebody means trusting him or her. Up to a point we can claim that every exercise of communication involves believing that our message will reach the recipient correctly and that it will not be misunderstood, or confidence based on some "charity principle" that other persons will not be trying to fool me all the time, confidence that is not present in people afflicted with paranoid beliefs. Of course, that requirement, which is basic in everyday communication, becomes much more stringent when the interactions move toward a greater intensity, as in family, friendship, and close-knit groups.

Discussion: Beliefs and normativity

The final points direct us toward a very sensitive issue: use and misuse of beliefs, and how to order them, or, rather, how to prevent abuses. This is a growing threat in a context dominated by new social media, with a huge flow of information, and where it becomes harder to assess which contents we can trust, in the midst of so much fake news. Believing becomes not merely a spontaneous activity, but a discipline that needs to be formed and to be built on a surer and more reliable ground. Such education programs would be aimed, for instance, to prevent predominant biases, like prestige and confirmation biases. A normative dimension derives from such awareness, a kind of "ethics of believing" should be assumed as a necessary chapter in the study and application of believing, a field that now receives more attention (Peels, 2016; Schmidt and Ernst, 2020).

The last application of believing is perhaps the most obvious, but not less subject to deep study and attention. Indeed, more analysis points to a convergence between religions and systems of meaning all placed under the umbrella of general belief systems, or sets of values, expectations, or faiths, able to provide meaning and purpose. These appear as a special kind of beliefs, with their own specific formation processes and characteristics, with central functions, and–again–unavoidable and not assimilated to other cognitive forms, like scientific scrutiny or sense perception.

All that has so far been discussed points to the importance of better studying and understanding the process of believing, often a pending issue in many areas, perhaps because of the dominance of reductionist models of cognition, which have neglected other forms as derived or secondary. We need good science to approach beliefs too, and to know better how they work, but not to replace them, something we could not, in any case, afford to do without a great anthropological and social cost.

Author contributions

The author confirms being the sole contributor of this work and has approved it for publication.

Funding

This manuscript was funded by Rüdiger Seitz, *via* the Volkswagen Foundation, Siemens Healthineers, and the Betz Foundation.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Angel, H. F., Oviedo, L., Paloutzian, R. F., Runehov, A. L. C., and Seitz, R. J. (2017). "Structuring credition," in *Processes of Believing: The Acquisition, Maintenance, and Change in Creditions*, eds H. F. Angel, L. Oviedo, R. F. Paloutzian, A. L. C. Runehov, and R. J. Seitz (Dordrecht: Springer International Publishing AG), 453–460. doi: 10.1007/978-3-319-50924-2_33

Castillo, R. D., Kloos, H., Richardson, M. J., and Waltzer, T. (2015). Beliefs as Self-Sustaining Networks: drawing Parallels between Networks of Ecosystems and Adults' Predictions. *Front. Psychol.* 6:1723. doi: 10.3389/fpsyg.2015.01723

Connors, M. H., and Halligan, P. W. (2015). A cognitive account of belief: a tentative road map. *Front. Psychol.* 5:1588. doi: 10.3389/fpsyg.2014.01588

Fuentes, A. (2019). *Why We Believe: Evolution and the Human Way of Being*. London: Yale University Press. doi: 10.12987/9780300249255

Miller, F. D. (2013). "Aristotle on Belief and Knowledge," in *Reason and Analysis in Ancient Greek Philosophy. Philosophical Studies Series*, eds G. Anagnostopoulos and F. Miller Jr. (Dordrecht: Springer), doi: 10.1007/978-94-007-6004-2_16

Peels, R. (2016). *Responsible Belief: A Theory in Ethics and Epistemology*. Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780190608118.001.0001

Schmidt, S., and Ernst, G. (2020). *The Ethics of Belief and Beyond: Understanding Mental Normativity*. London: Routledge. doi: 10.4324/9780429284977

Shagan, E. H. (2021). *The Birth of Modern Belief: Faith and Judgement from the Middle Ages to the Enlightenment*. Princeton: Princeton University Press.

Shermer, M. (2011). *The Believing Brain: From Ghosts and Gods to Politics and Conspiracies — How We Construct Beliefs and Reinforce Them as Truths*. London: MacMillan.

Sunstein, C. (2021). *This is Not Normal: The Politics of Everyday Expectation*. Yale: Yale University Press.

frontiers | Frontiers in Psychology

OPINION
published: 19 May 2022
doi: 10.3389/fpsyg.2022.913603

Believing and Appraising in Context: Cognizing Experiences as Events

Ann Taves[1]* *and Raymond F. Paloutzian*[2]

[1] Department of Religious Studies, University of California, Santa Barbara, Santa Barbara, CA, United States, [2] Department of Psychology, Westmont College, Santa Barbara, CA, United States

Keywords: belief, believing, credition, event, event cognition, meaning making

It is clear that schemas and expectations unconsciously affect how people interpret their experience in real time (van Elk and Aleman, 2017; Van Leeuwen and van Elk, 2019). It is also clear that cultural beliefs as manifest in worldviews and ways of life affect how people consciously reflect on and reappraise their experiences (Kelly, 1955; Saroglou and Cohen, 2011, 2013). How these two processes are linked is not so clear. Here we argue that event cognition not only supplies a promising bridge between unconscious and conscious information processing, but allows us to further integrate research on unconscious appraisal processes (Scherer et al., 2001), conscious attributions (Kelley and Michela, 1980; Malle, 1999, 2004), and multi-level approaches to meaning making (Park and Folkman, 1997; Park, 2010; Markman et al., 2013) and believing (for an overview, see Seitz and Angel, 2020)[1].

OPEN ACCESS

Edited by:
Douglas A. MacDonald, University of Detroit Mercy, United States

Reviewed by:
Harris Friedman, University of Florida, United States

***Correspondence:**
Ann Taves
anntaves@ucsb.edu

Specialty section:
This article was submitted to Theoretical and Philosophical Psychology, a section of the journal Frontiers in Psychology

Received: *05 April 2022*
Accepted: *02 May 2022*
Published: *19 May 2022*

Citation:
Taves A and Paloutzian RF (2022) Believing and Appraising in Context: Cognizing Experiences as Events. Front. Psychol. 13:913603. doi: 10.3389/fpsyg.2022.913603

WHAT IS AN EVENT?

The *Oxford English Dictionary* defines *events* simply as things that happen. Philosophers tend to begin with this general definition when discussing events, noting that "this definition merely shifts the burden [of definition] to clarifying the meaning of 'happen'" (Casati and Varzi, 2020). For scientific purposes, it is important to distinguish between things that happen regardless of whether there is a living organism to perceive them and happenings that living organisms perceive. Molecules move and stars collide regardless of whether living organisms are present. In doing so happenings generate information. Because far more information is generated than any organism can perceive and process, organisms have evolved to perceive and process the information that they need to survive. Humans and some other animals do so by segmenting information, that is, by a cognitive process that divides it into chunks with a beginning and an end (Baldwin and Kosie, 2021; Ross and Easton, 2022).

Researchers who study this cognitive process define an event more technically as "a segment of time at a given location conceived by an observer to have a beginning and an end" (Zacks and Tversky, 2001, p. 3). This means that an event is spatially and temporally located and is perceived from the point of view of the subject. An event is constructed as the brain chunks the dynamic flow of incoming information into segments. Most researchers also agree that the extracted segment is perceived as coherent and causally related, which points to the basis on which we chunk the flow of information (Hohwy et al., 2021).

[1] Psychological interest in events emerged within both clinical and experimental psychology to produce lines of research that focus on differences levels of analysis. In clinical psychology, Kelly's (1955) personal construct theory, which is centrally concerned with the role of constructs in the interpretation of events, was influenced by hermeneutic approaches in the humanities (Butt and Warren, 2016). In experimental psychology, the focus has been on the perceptual processes (i.e., mechanisms) involved in cognizing events (Shipley, 2008; Zacks, 2020).

The cognitive definition of an event has implications for how we think about experience (an uncountable noun) and experiences (a plural noun). The former designates the flow of information of which we are aware. The latter are self-reported events that have a beginning and an end; this means that from a cognitive perspective, experiences are events regardless of whether they are externally verifiable or not (for fuller discussion, see Taves and Asprem, 2017). Treating experiences as events allows us to investigate how the flow of information of which we are aware (experience) is transformed cognitively into structured units (experiences) with a beginning and end that we can describe, remember, and recount if and when we attend to them. Doing so allows us to link subpersonal and personal levels of analysis and integrate several lines of research in psychology, including multi-level approaches to meaning making and believing.

What then is not an event? Although philosophers debate this issue, they generally agree that physical objects in isolation are not events and that events occur when things change or interact (Shipley, 2008; Casati and Varzi, 2020). If we extend the criterion of change to mental things, such as ideas, beliefs, concepts, and goals, then beliefs all by themselves are not events. Events occur when beliefs change or interact, i.e., in the process of believing. How then do we cognize events?

WORKING MODELS AS PROBABILISTIC REPRESENTATIONS

Most researchers understand the processes by which events are initially cognized within a predictive processing framework (Hohwy et al., 2021).

Bottom-up input is weighed against a top-down prediction of what is happening based on prior experience. The prediction of what is happening is represented in a working model of an event. The working model is, thus, is a probabilistic assessment (i.e., appraisal) of incoming information in light of prior experience. The model is retained as long as it more or less fits with the incoming information. If there is a significant change in the input, an error signal is generated, which leads to a new or revised event model (Radvansky and Zacks, 2014, 2017; Zacks, 2020). The process of assessing the incoming information can be viewed as a meaning making process and the probabilistic assessment as an appraisal. The working model that is generated based on this assessment is a representation of the event. It can also be viewed as an implicit belief regarding what happened (Seitz et al., 2017, 2018; Paloutzian et al., 2021).

The working model is generated in working memory and thus is fleeting (like the dream you can't remember) unless transferred to long-term memory (Zacks, 2020, pp. 172–177). If the event is stored in long-term memory, we can remember it, narrate it, and reflect on it. In other words, the working model links what we consciously experience as happening with underlying cognitive processes and, if retained in long-term memory, allows us to recall and reflect on past events. Each time we recall an event, we construct a new event model in working memory with its own new spatio-temporal context.

In sum, the sensory input from the body, environment, and prior experience interact to form a working model of what is happening. Insofar as we are conscious of the contents of working memory, we are conscious of the contents of the working model. That is how we experience the event and come to believe—at least implicitly—that an event occurred (Seitz et al., 2022). If our experience is stored in long-term memory, we can remember and recount it. But these are separate events with their own working models of what is happening.

WORKING MODELS AND PRIOR EXPERIENCE

The working model is based on a probabilistic assessment of incoming information *in light of prior experience* (Zacks, 2020, pp. 177–180). In psychological terms, the brain assesses incoming information in light of schemas and expectations. Schemas and expectations, although likely built on shared, reliably developing templates, are typically elaborated in culturally specific ways and acquired through cultural learning. Each component of a working model—time, space, objects, sensations, relations, and causes—draws on prior experience. With respect to time, schemas provide an expected time frame for the event. With respect to space, schemas identify the specific place or type of space in which the event is occurring. Schemas and expectations determine the types of agents—visible and invisible—that may be involved, identify who is involved, what they are doing and why. They allow us to recognize the objects involved and assess what is happening to them. They allow us to identify what we are sensing or feeling. Finally, overall event schemas help us understand how all these things are related and what is causing it to happen. Because working models rely on prior experience, we would expect infants, foreigners, and experts to have representations of an event that differ from those of the average culturally literate adult.

EVOLVED AND CULTURAL KNOWLEDGE

Although people's representations of an event differ based on their prior knowledge, human's evolved capacity to chunk the flow of information into events means that everyone—including infants—can recognize that something has happened. Thus, there is growing evidence that infants can attend to structured patterns in goal-directed activity and that these patterns provide a basic sense of where to segment the dynamic flow (Levine et al., 2019; Zacks, 2020). When we confront a new situation, we draw on those basic capabilities. Then, as we grow and develop in a particular time and place, we become able to comprehend and recall events with greater ease and accuracy. We become more "fluent." We learn to pick out the relevant details of events that allow us to efficiently predict what is happening and guide us in deciding what to do. We do so in the context of particular culturally distinct ways of life. The event models of fluent adults are culture specific; they include appraisals and beliefs guided by and adapted to the culture in which one is imbedded (Baldwin and Kosie, 2021).

To illustrate, imagine a cultural event such as going to a Catholic Mass. A young child would understand that the event had a beginning and an end and contained any number of subevents, such as entering and leaving the church, walking to the front the church with their parents, and returning to their seats. When a bell was rung, the child's attention might be drawn to the altar and the man who is doing something there. An adult who knew nothing of Catholic ritual would also recognize these and other subevents. If they had experienced other rituals, they would likely recognize it as such, without knowing much else. Culturally fluent Catholics, however, would have internalized an event model of the Mass as a ritual that recapitulates Jesus' death and resurrection and his promise to be present in the sharing of bread and wine. They would know that the words of the priest over the bread and wine make Christ present and, thus, why the bell draws their attention to that point in the service.

RE-APPRAISING EVENTS

Events, as they initially surface to awareness, can take many forms. Most are routine; we give them little thought. Some events, however, stand out because they are puzzling, surprising, disturbing, or life-changing. These are the events we tend to remember, recount, and in some cases reappraise. If an experience does not fit with what we have learned to expect or believe, we return to it to try to figure out what happened or how to cope with what we know or believe happened. This process of making meaning out of ambiguity, appraising it, remaking meaning, reappraisal, and so on has been well researched and documented as a series of events (Park and Folkman, 1997; Park, 2010; Markman et al., 2013).

Understanding the meaning making process as a series of events, therefore, implies that the initial appraisal of meaning takes place as part of the initial event. This generates what Park (2010) refers to as the "situational" meaning of the event. But when someone consciously assesses their initial sense of what happened in light of their overall set of beliefs and goals, that constitutes a subsequent event, and the processes of coping with discrepancies between the situational and global meanings generates a whole series of additional events with their modified beliefs.

The working model of what is happening now, thus, allows us to think in two directions. We can (1) think about change over time as an initial event that is reappraised in the context of subsequent events, or we can (2) think about the levels of processing that give rise to an event model in the context of a single event. The first is an explicit reflective cultural process; the second relies on culturally learned expectations and schemas that function as priors in the probabilistic assessment of what is happening at any given moment. Treating experiences as events allows us to consider the components that interact to generate an experience and compare the interaction of schemas and expectations with phenomenological features of experiences in a variety of different cultural contexts (for a full discussion, see Taves and Barlev, 2022).

AUTHOR CONTRIBUTIONS

AT conceptualized the concept of events and how minds cognize events as related to the chunking and perception of things happening in real time. RP added their relationship to the processes of believing. Both authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

We are grateful to Elliott Ihm, Rüdiger Seitz, and Hans-Ferdinand Angel for feedback on earlier drafts of this manuscript.

REFERENCES

Baldwin, D. A., and Kosie, J. E. (2021). How does the mind render streaming experience as events? *Top. Cogn. Sci.* 13, 79–105. doi: 10.1111/tops.12502

Butt, T., and Warren, B. (2016). "Personal construct theory and philosophy." in *The Wiley Handbook of Personal Construct Psychology,* eds. D. A. Winter and N. Reed. New York: John Wiley and Sons. doi: 10.1002/9781118508275.ch2

Casati, R., and Varzi, A. (2020). "Events." in *Stanford Encyclopedia of Philosophy,* ed. E. N. Zalta. Available online at: https://plato.stanford.edu/archives/sum2020/entries/events/

Hohwy, J., Hebblewhite, A., and Drummond, T. (2021). Events, event prediction, and predictive processing. *Top. Cogn. Sci.* 13, 252–255. doi: 10.1111/tops.12491

Kelley, H. H., and Michela, J. L. (1980). Attribution theory and research. *Ann. Rev. Psychol.* 31, 457–501. doi: 10.1146/annurev.ps.31.020180.002325

Kelly, G. A. (1955). *Psychology of Personal Constructs: Theory and personality*. New York: Norton.

Levine, D., Buchsbaum, D., Hirsh-Pasek, K., and Golinkoff, R. M. (2019). Finding events in a continuous world: a developmental account. *Develop. Psychobiol.* 61, 376–389. doi: 10.1002/dev.21804

Malle, B. F. (1999). How people explain behavior: a new theoretical framework. *Personal. Social Psychol. Rev.* 3, 23–48. doi: 10.1207/s15327957pspr0301_2

Malle, B. F. (2004). *How the Mind Explains Behavior: Folk Explanations, Meaning, and Social Interaction*. Cambridge, MA: MIT Press. doi: 10.7551/mitpress/3586.001.0001

Markman, K. D., Proulx, T., and Lindberg, M. J. (Eds.) (2013). *The Psychology of Meaning*. Washington, D.C.: American Psychological Association. doi: 10.1037/14040-000

Paloutzian, R. F., Seitz, R. J., and Angel, H.-F. (2021). "The processes of believing and communicating with the unseen." in *Human interaction with the divine, the sacred, and the deceased: Psychological, scientific, and theological perspectives,* eds. T. G. Plante and G. E. Schwartz (1st ed.) (New York: Routledge), 213–233. doi: 10.4324/9781003105749-14

Park, C. L. (2010). Making sense of the meaning literature: an integrative review of meaning making and its effects on adjustment to stressful life events. *Psychol Bull.* 136, 257–301. doi: 10.1037/a0018301

Park, C. L., and Folkman, S. (1997). Meaning in the context of stress and coping. *Rev. General Psychol.* 1, 115–144. doi: 10.1037/1089-2680.1.2.115

Radvansky, G. A., and Zacks, J. M. (2014). *Event cognition*. New York: Oxford University Press. doi: 10.1093/acprof:oso/9780199898138.001.0001

Radvansky, G. A., and Zacks, J. M. (2017). Event boundaries in memory and cognition. *Curr Opin Behav. Sci.* 17, 133–140. doi: 10.1016/j.cobeha.2017.08.006

Ross, T. W., and Easton, A. (2022). The Hippocampal Horizon: Constructing and Segmenting Experience for Episodic Memory. *Neurosci. Biobehav. Rev.* 132, 181–196. doi: 10.1016/j.neubiorev.2021.11.038

Saroglou, V., and Cohen, A. B. (2011). Psychology of culture and religion: introduction to the *JCCP* special issue. *J. Cross Cult. Psychol.* 42, 1309–1319. doi: 10.1177/0022022111412254

Saroglou, V., and Cohen, A. B. (2013). "Cultural and cross-cultural psychology of religion." in *Handbook of the Psychology of Religion and Spirituality,* eds. R. F. Paloutzian and C. L. Park (Second Edition) (New York: Guilford Press), 330–354.

Scherer, K. R., Schorr, A., and Johnstone, T. (eds.) (2001). *Appraisal Processes in Emotion: Theory, Methods, Research* (1 edition) (New York: Oxford University Press).

Seitz, R. J., and Angel, H.-F. (2020). Belief formation—A driving force for brain evolution. *Brain Cogn.* 140, 105548. doi: 10.1016/j.bandc.2020.105548

Seitz, R. J., Angel, H.-F., and Paloutzian, R. F. (2022). *Bridging the Gap Between Believing and Memory Functions*.

Seitz, R. J., Paloutzian, R. F., and Angel, H.-F. (2017). Processes of believing: where do they come from? What are they good for? *F1000Research*. *5*, 2573. doi: 10.12688/f1000research.9773.2

Seitz, R. J., Paloutzian, R. F., and Angel, H.-F. (2018). From believing to belief: a general theoretical model. *J. Cogn. Neurosci.* 30, 1254–1264. doi: 10.1162/jocn_a_01292

Shipley, T. F. (2008). "An invitation to an event." in *Understanding Events: From Perception to Action,* T. eds. F. Shipley and J. M. Zacks (New York: Oxford University Press), 3–30.

Taves, A., and Asprem, E. (2017). Experience as event: event cognition and the study of (religious) experiences. *Religion, Brain, and Behavior* 7, 43–62. doi: 10.1080/2153599X.2016.1150327

Taves, A., and Barlev, M. (2022). A Feature-Based Approach to the Comparative Study of "Nonordinary" Experiences. *American Psychologist*. doi: 10.1037/amp0000990

van Elk, M., and Aleman, A. (2017). Brain mechanisms in religion and spirituality: an integrative predictive processing framework. *Neurosci Biobehav Rev*, *73*, 359–378. doi: 10.1016/j.neubiorev.2016.12.031

Van Leeuwen, N., and van Elk, M. (2019). Seeking the supernatural: the interactive religious experience model. *Religion Brain Behav.* 9, 221–251. doi: 10.1080/2153599X.2018.1453529

Zacks, J. M. (2020). Event perception and memory. *Ann. Rev. Psychol.* 71, 165–191. doi: 10.1146/annurev-psych-010419-051101

Zacks, J. M., and Tversky, B. (2001). Event structure in perception and conception. *Psychol. Bull.* 127, 3–21. doi: 10.1037/0033-2909.127.1.3

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Taves and Paloutzian. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

OPINION
published: 23 March 2022
doi: 10.3389/fpsyg.2022.844891

Meaning Making Following Trauma

*Crystal L. Park**

Department of Psychological Sciences, University of Connecticut, Storrs, CT, United States

Keywords: meaning, appraisal, cognitive processing, beliefs, goals

INTRODUCTION

The meaning making model provides a useful framework for integrating myriad existing meaning-related theories and empirical findings. In this overview, I describe this model, which comprises both global and situational (event-specific) aspects. Global meaning encompasses foundational beliefs, values and goals, and a subjective sense of meaningfulness while situational meaning entails the appraisal of an experience. When an experience is perceived as discrepant with global meaning, individuals experience distress and engage in a variety of efforts to make meaning of that experience. Meaning making is usually aimed at changing the meaning of the situation but can also involve changing global meaning (e.g., adopting a new way of understanding the world or new goals; i.e., *meaning made*). Successful meaning making reduces discrepancies between global meaning and individuals' assigned meaning of the specific experience and restores harmony within their global meaning vis-à-vis their current experience.

The model of meaning making described here is based on a growing body of research regarding responses to adversity, such as serious illness, bereavement, sexual assault, incest, the COVID-19 pandemic, natural disasters, and terrorist attacks [Fitzke et al., 2021; see Park et al. (2017) and Park and Blake (2020), for reviews]. This model distinguishes two levels of meaning: global (people's fundamental and overarching beliefs and their hierarchies of goals and values; Park, 2010) and situational (how global meaning, in conjunction with a given particular context, influences assigning meaning and responding to a particular situation; Park, 2017; **Figure 1**).

GLOBAL MEANING

Global meaning refers to individuals' foundational orienting systems (Trevino et al., 2019), consisting of individuals' deeply held beliefs regarding reality, such as fairness, control, and identity (Park, 2017; Clifton et al., 2019; Pilkington et al., 2021) as well as their goals (states that one desires and pursues or that one already possesses and seeks to maintain, such as health, wealth, or family relationships; Lewis, 2020) In addition, global meaning includes a subjective sense of life as meaningful (e.g., purposeful, comprehensible; Park, 2010).

SITUATIONAL MEANING

In addition to global meaning systems, psychological adjustment is influenced by one's circumstances and how those circumstances are understood (i.e., their situational meaning). People continuously monitor their experiences and assign meaning to (i.e., appraise) them. Encountering potentially difficult or stressful situations leads to determining the extent to which it is discrepant with

OPEN ACCESS

Edited by:
Rüdiger J. Seitz, Heinrich Heine University of Düsseldorf, Germany

Reviewed by:
Nina Dalkner, Medical University of Graz, Austria
Gary D. Bond, Eastern New Mexico University, United States

***Correspondence:**
Crystal L. Park
crystal.park@uconn.edu

Specialty section:
This article was submitted to Theoretical and Philosophical Psychology, a section of the journal Frontiers in Psychology

Received: *28 December 2021*
Accepted: *27 January 2022*
Published: *23 March 2022*

Citation:
Park CL (2022) Meaning Making Following Trauma. Front. Psychol. 13:844891. doi: 10.3389/fpsyg.2022.844891

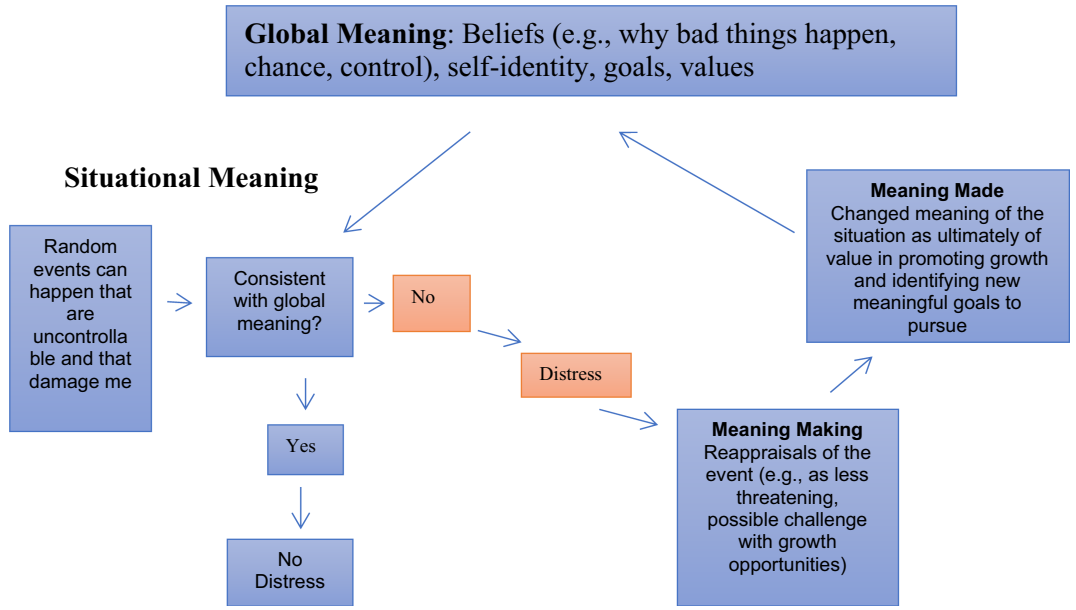


FIGURE 1 | Meaning making model as applied to a negative situation.

one's global meaning, and to the extent it is, coping with and making meaning of those experiences and adjusting to them.

Appraisals of Events

People appraise, or assign a particular meaning, to their encounters, determining the extent to which they are threatening and controllable, attributing causes, and discerning their implications (Park, 2010). These appraised meanings, in turn, determine individuals' reactions to those events. Highly traumatic events are commonly appraised as unpredictable, unfair, and uncontrollable and as having pervasive adverse implications for survivors and their futures (Brown et al., 2019). The meaning making model asserts that distress is not generated by the appraised meaning itself but rather by discrepancies between that appraised meaning and the individual's global meaning system (Park, 2010; Park et al., 2016). For example, a study of pregnant women who experienced the Queensland Flood found that appraising the consequences of the flood on themselves and their families predicted later depression and anxiety symptoms, but appraising the consequences as positive buffered the long-term effects of peritraumatic distress on anxiety levels in these new mothers 2 years later (Paquin et al., 2021).

Appraised Violations of Global Meaning

After people appraise or assign meaning to an event, they determine the degree to which it is consistent or discrepant with their global meaning. Perceived discrepancies (e.g., with their sense that the world is understandable and fair or that the event is not what they wanted to have happened) produce distress (Steger et al., 2015; Park et al., 2016). A scale to assess this global meaning violation was developed recently, the Global Meaning Violations Scale (GMVS; Park et al., 2016). A study of college students reporting on their most stressful event using the GMVS demonstrated that violations of global beliefs and violations of global goals were each independently related to distress. Similarly, a recent study of a national sample early in the COVID-19 pandemic showed that greater belief violations were associated with higher levels of anxiety and depression symptoms (Milman et al., 2020).

Meaning Making

This violation-related distress is painful, motivating people to try to alleviate it. These efforts can involve meaning making, although people also engage in many other strategies to try to reduce their distress, including a variety of active and passive coping strategies (e.g., Park et al., 2021). Meaning making aims to restore disrupted global meaning through approach-oriented intrapsychic attempts to develop new and acceptable ways of understanding the situation that are more consistent with one's global meaning or by changing one's global meaning beliefs and goals. Following successful meaning making, people have a different view of the situation and have modified their beliefs and goals to regain consistency among them, an outcome termed *meaning made* (Park, 2017).

Severe trauma can disrupt a person's global meaning (Janoff-Bulman, 1989). Making meaning typically involves cognitive processing of appraised and global meanings to change or reframe them and make them more consistent (Fitzke et al., 2021; Huang et al., 2021). People can change situational appraisals to better integrate them into their global meaning system (assimilation), such as coming to see the event as less damaging or, perhaps, even positive in its consequences (Paquin et al., 2021; Park and Boals, 2021). For appraisals of events that are highly discrepant with global meaning, meaning making may

require changing one's global meaning to accommodate the trauma.

Positive reappraisal, looking for ways to view the situation in a more positive light, or focusing on identifying positive attributes of an event and reminding oneself of those benefits, are very common meaning-making strategies (Park, 2010). A longitudinal study of a national sample of Americans (the Midlife in the United States Series Study) found that attempting to making meaning of a highly stressful or traumatic life event through positive reappraisal was associated with subsequent higher levels of positive mood and lower levels of negative mood years later (Fitzke et al., 2021). Other meaning making strategies include seeking more benign explanations for the situation and making downward comparisons with real or hypothetical others in relatively poorer straits (Gerber et al., 2018). Several studies of individuals dealing with the COVID-19 pandemic have highlighted some of the different strategies people use in efforts at making meaning of the pandemic, including through positive reinterpretation (Park et al., 2021), seeking out potential benefits or growth (Yang et al., 2021), and accepting the reality of the situation (Umucu and Lee, 2020).

Meanings Made

Meaning making processes can be helpful by making new meanings, that is, changes in appraised or global meaning resulting from the cognitive processing involved in meaning making. Sometimes individuals change their understanding of the reason the event occurred, developing a more benign understanding; this new and more benign view is a key type of situational meaning made (e.g., Beierl et al., 2020). People may also perceive that they have changed in positive ways as a result of the trauma, such as improved relationships or enhanced coping skills (e.g., Park and Boals, 2021). Global beliefs and goals can change as well. For example, a study of Norwegian adults who survived the Southeast Asian tsunami in 2004 found those who reported their beliefs about the world changed in a positive way experienced fewer posttraumatic stress disorder (PTSD) symptoms and better quality of life (Nygaard and Heir, 2012). Survivors may change the goals they pursue as well, letting go of goals that are no longer realistic and doubling down on more attainable goals (Haase et al., 2021).

Discrepancy Reduction Leads to Better Adjustment (Through Meanings Made)

People make meaning as a way to reduce discrepancies between situational and global meanings, and greater reductions in the size of discrepancies predicts better adjustment following trauma. For example, in a study of college students reporting on their most stressful or traumatic life event, reductions in global meaning violations over time was associated with concomitant reductions in PTSD symptoms (Park et al., 2016). Similarly, in a study of military veterans, reductions in Self-Blame through cognitive processing therapy, which relies strongly on meaning-making, were associated with reductions in PTSD symptoms (Holliday et al., 2018). On the other hand, continued inability to integrate one's appraisal of a traumatic event into global meaning often leads to continued rumination, intrusive thoughts, and depression (Zakarian et al., 2019; Huang et al., 2021).

To date, findings from research conducted in many different trauma and stress contexts supports linkages among components of the meaning-making model. However, few studies have fully examined the set of linkages outlined in the meaning making model. For example, few studies have assessed violations of beliefs and goals nor whether meaning making efforts following trauma help make meaning by reducing discrepancies between appraised situational and global meaning. To adequately study these linkages would require longitudinal studies assessing initial situational appraisals, violations and distress, meaning making efforts, meanings made, changes in violations and subsequent adjustment. Instead, most of the work on meaning making is cross-sectional and retrospective (e.g., Huang et al., 2021), with only a few multiple-time point studies examining these issues (e.g., Fitzke et al., 2021). While this growing body of work suggests that discrepancy reductions mediate effects of meaning making and meanings made on adjustment, much remains to be learned about meaning making and its relations to managing and overcoming trauma.

CONCLUSION

The meaning making model is a useful framework for integrating existing meaning-related theories and empirical findings (Park, 2010). To date, however, research has tested the model in piecemeal fashion, focusing on specific components or linkages, and findings are essentially supportive of the meaning making model. More inclusive longitudinal research focusing on relationships among various components as people engage in meaning making of and adjust to highly stressful situations will provide more rigorous tests of the model. Ultimately, the meaning making model will provide insight and clinical applications (Park et al., 2017) to promote better adjustment to highly stressful experiences.

AUTHOR CONTRIBUTIONS

CLP developed the ideas contained in this paper, conducted the literature review, and wrote and edited the manuscript.

REFERENCES

Beierl, E. T., Böllinghaus, I., Clark, D. M., Glucksman, E., and Ehlers, A. (2020). Cognitive paths from trauma to posttraumatic stress disorder: a prospective study of Ehlers and Clark's model in survivors of assaults or road traffic collisions. *Psychol. Med.* 50, 2172–2181. doi: 10.1017/S0033291719002253

Brown, L. A., Belli, G. M., Asnaani, A., and Foa, E. B. (2019). A review of the role of negative cognitions about oneself, others, and the world in the treatment of PTSD. *Cogn. Ther. Res.* 43, 143–173. doi: 10.1007/s10608-018-9938-1

Clifton, J. D. W., Baker, J. D., Park, C. L., Yaden, D. B., Clifton, A. B. W., and Seligman, M. E. P. (2019). Primal world beliefs. *Psychol. Assess.* 31, 82–99. doi: 10.1037/pas0000639

Fitzke, R. E., Marsh, D. R., and Prince, M. A. (2021). A longitudinal investigation of the meaning-making model in midlife adults who have experienced trauma. *J. Clin. Psychol.* 77, 2878–2893. doi: 10.1002/jclp.23272

Gerber, J. P., Wheeler, L., and Suls, J. (2018). A social comparison theory meta-analysis 60+ years on. *Psychol. Bull.* 144, 177–197. doi: 10.1037/bul0000127

Haase, C. M., Singer, T., Silbereisen, R. K., Heckhausen, J., and Wrosch, C. (2021). Well-being as a resource for goal reengagement: evidence from two longitudinal studies. *Motiv. Sci.* 7, 21–31. doi: 10.1037/mot0000199

Holliday, R., Holder, N., and Surís, A. (2018). Reductions in self-blame cognitions predict PTSD improvements with cognitive processing therapy for military sexual trauma-related PTSD. *Psychiatry Res.* 263, 181–184. doi: 10.1016/j.psychres.2018.03.007

Huang, M., Schmiedek, F., and Habermas, T. (2021). Only some attempts at meaning making are successful: the role of change-relatedness and positive implications for the self. *J. Pers.* 89, 175–187. doi: 10.1111/jopy.12573

Janoff-Bulman, R. (1989). Assumptive worlds and the stress of traumatic events: applications of the schema construct. *Soc. Cogn.* 7, 113–136. doi: 10.1521/soco.1989.7.2.113

Lewis, N. A. (2020). Purpose in life as a guiding framework for goal engagement and motivation. *Soc. Personal. Psychol. Compass* 14, 1–11. doi: 10.1111/spc3.12567

Milman, E., Lee, S. A., Neimeyer, R. A., Mathis, A. A., and Jobe, M. C. (2020). Modeling pandemic depression and anxiety: the mediational role of core beliefs and meaning making. *J. Affect. Disord. Rep.* 2:100023. doi: 10.1016/j.jadr.2020.100023

Nygaard, E., and Heir, T. (2012). World assumptions, posttraumatic stress and quality of life after a natural disaster: a longitudinal study. *Health Qual. Life Outcomes* 10:76. doi: 10.1186/1477-7525-10-76

Paquin, V., Elgbeili, G., Laplante, D. P., Kildea, S., and King, S. (2021). Positive cognitive appraisal "buffers" the long-term effect of peritraumatic distress on maternal anxiety: the Queensland flood study. *J. Affect. Disord.* 278, 5–12. doi: 10.1016/j.jad.2020.09.041

Park, C. L. (2010). Making sense of the meaning literature: an integrative review of meaning making and its effects on adjustment to stressful life events. *Psychol. Bull.* 136, 257–301. doi: 10.1037/a0018301

Park, C. L. (2017). Distinctions to promote an integrated perspective on meaning: global meaning and meaning-making processes. *J. Constr. Psychol.* 30, 14–19. doi: 10.1080/10720537.2015.1119082

Park, C. L., and Blake, E. B. (2020). "Resilience and recovery following disasters: the meaning making model," in *Positive Psychological Approaches to Disaster: Meaning, Resilience, and Posttraumatic Growth*. ed. S. E. Schulenberg (Cham: Springer), 9–25.

Park, C. L., and Boals, A. (2021). "Current assessment and interpretation of perceived post-traumatic growth," in *Redesigning Research on Post-Traumatic Growth*. eds. F. J. Infurna and E. Jayawickreme (Oxford: Oxford University Press), 12–27.

Park, C. L., Currier, J. M., Harris, J. I., and Slattery, J. M. (2017). *Trauma, Meaning, and Spirituality: Translating Research into Clinical Practice*. Washington, DC: American Psychological Association.

Park, C. L., Finkelstein-Fox, L., Russell, B. S., Fendrich, M., Hutchison, M., and Becker, J. (2021). Psychological resilience early in the COVID-19 pandemic: stressors, resources, and coping strategies in a national sample of Americans. *Am. Psychol.* 76, 715–728. doi: 10.1037/amp0000813

Park, C. L., Riley, K. E., George, L. S., Gutierrez, I. A., Hale, A. E., Cho, D., et al. (2016). Assessing disruptions in meaning: development of the global meaning violation scale. *Cogn. Ther. Res.* 40, 831–846. doi: 10.1007/s10608-016-9794-9

Pilkington, P. D., Bishop, A., and Younan, R. (2021). Adverse childhood experiences and early maladaptive schemas in adulthood: a systematic review and meta-analysis. *Clin. Psychol. Psychother.* 28, 569–584. doi: 10.1002/cpp.2533

Steger, M. F., Owens, G. P., and Park, C. L. (2015). Violations of war: testing the meaning-making model among Vietnam veterans. *J. Clin. Psychol.* 71, 105–116. doi: 10.1002/jclp.22121

Trevino, K. M., Pargament, K. I., Krause, N., Ironson, G., and Hill, P. (2019). Stressful events and religious/spiritual struggle: moderating effects of the general orienting system. *Psychol. Relig. Spiritual.* 11, 214–224. doi: 10.1037/rel0000149

Umucu, E., and Lee, B. (2020). Examining the impact of COVID-19 on stress and coping strategies in individuals with disabilities and chronic conditions. *Rehabil. Psychol.* 65, 193–198. doi: 10.1037/rep0000328

Yang, Z., Ji, L. J., Yang, Y., Wang, Y., Zhu, L., and Cai, H. (2021). Meaning making helps cope with COVID-19: a longitudinal study. *Personal. Individ. Differ.* 174:110670. doi: 10.1016/j.paid.2021.110670

Zakarian, R. J., McDevitt-Murphy, M. E., Bellet, B. W., Neimeyer, R. A., and Burke, L. A. (2019). Relations among meaning making, PTSD, and complicated grief following homicide loss. *J. Loss Trauma* 24, 279–291. doi: 10.1080/15325024.2019.1565111

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Park. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Behavioral Neuroscience

OPINION
published: 05 July 2022
doi: 10.3389/fnbeh.2022.929332

Gut Signals and Gut Feelings: Science at the Interface of Data and Beliefs

*Peter Holzer**

Division of Pharmacology, Otto Loewi Research Centre, Medical University of Graz, Graz, Austria

Keywords: brain-gut axis, gut-brain axis, interoception, emotion, affect, cognition, beliefs

INTRODUCTION

The discovery that signals from the gut reaching the brain can have an impact on affect, emotion and cognition including beliefs and decision-making has been met with considerable public attention. As discussed here, widespread interest in a research field that touches one's personal health also raises expectations and beliefs that are not thoroughly backed by validated scientific evidence. The term "gut feelings," derived from a lay interpretation of the impact of gut signals on mental activity, is a popular but scientifically ill-defined term that may even lead science astray from its principles of investigation. In these interdependencies, the relation between gut signals and "gut feelings" is a worthwhile theme for analysis in credition research. As outlined in the overview of the Special Topic that this article is part of, credition research is an interdisciplinary approach to understand the nature of beliefs and believing. Based on distinct neuronal processes in the brain, credition refers to the integrative processing of information, its valuation in terms of personal meaning and its impact on a person's behavioral decisions (Seitz et al., 2018; Seitz, 2022).

BIDIRECTIONAL GUT-BRAIN COMMUNICATION

Psychosomatic medicine has gathered ample evidence that gastrointestinal function can be altered by emotions and stress. For instance, 75 years ago Almy and Tulin (1947) published a study in which they performed sigmoidoscopies in volunteer medical students. When during the examination they told the students that they had discovered a carcinoma, they instantly observed strong muscle contractions and an increase in blood flow in the rectum. Once they explained the hoax, the uproar in the rectum subsided rapidly. Although this kind of study would no longer receive ethical approval (Shea-Donohue et al., 2005), it shows that emotional stress can have an immediate impact on the gut. Since then, many studies have confirmed (for reviews see Mayer, 2000; Taché et al., 2001), in a more humane way, that acute physical and emotional stress can affect the digestive tract in a regionally distinct manner, retarding gastric emptying ("being heavy on the stomach") but hastening colonic propulsion ("soiling one's pants"). The changes in gut function accompanying long-term stress, however, may substantially differ from those in acute settings (McEwen, 2007). The communication from the brain to the gut, often referred to as "brain-gut axis," is transmitted by several pathways (**Figure 1**) including the efferent autonomic nervous system with its sympathetic and parasympathetic divisions and neuroendocrine factors of the sympathetic-adrenal medulla and hypothalamus-pituitary-adrenal cortex systems (Holzer et al., 2015).

OPEN ACCESS

Edited by:
Rüdiger J. Seitz, Heinrich Heine University of Düsseldorf, Germany

Reviewed by:
Hirokata Fukushima, Kansai University, Japan
Tian Yuan, Wake Forest School of Medicine, United States
Patrice Duquette, Private Practitioner, Birmingham, MI, United States

***Correspondence:**
Peter Holzer
peter.holzer@medunigraz.at

Specialty section:
This article was submitted to Emotion Regulation and Processing, a section of the journal Frontiers in Behavioral Neuroscience

Received: *26 April 2022*
Accepted: *01 June 2022*
Published: *05 July 2022*

Citation:
Holzer P (2022) Gut Signals and Gut Feelings: Science at the Interface of Data and Beliefs. Front. Behav. Neurosci. 16:929332. doi: 10.3389/fnbeh.2022.929332

Communication between brain and gut is in fact a bidirectional process that, besides efferent connections, also involves afferent pathways that carry information from the gastrointestinal tract to the central nervous system (CNS) (**Figure 1**). Extrinsic sensory neurons of the vagal and spinal nerves transmitting mechanical and chemical stimuli are one component of this "gut-brain axis." Endocrine chemical messages carried by gut hormones released from enteroendocrine cells in the gut mucosa, mediators (cytokines) of the gastrointestinal immune system, metabolic factors related to nutrient absorption and digestion, and messengers generated by the gut microbiota (**Figure 1**) constitute other important components of the gut-brain communication network. The enteric nervous system intrinsic to the gut (Perez-Burgos et al., 2014), enteroendocrine cells (Kaelberer et al., 2020) as well as immune and microbial messengers also use, in part, extrinsic sensory neurons to signal to the brain (Holzer et al., 2015).

This complex afferent communication system provides the brain with integrated information on gut function. In this task, the gut-brain axis contributes to interoception, a process that enables the brain to "know" the internal state of the body (Craig, 2002) and align its mental activity and homeostatic body control (Chen et al., 2021). Functional imaging studies have shown that information coming from the gut reaches brain regions relevant to emotion, affect and cognition. A good deal of what we now know about the bidirectional information exchange between gut and brain has been disclosed by research efforts to understand irritable bowel syndrome (IBS). Characterized by recurrent abdominal pain associated with alterations in bowel habits, IBS is frequently comorbid with anxiety disorders and depression and now widely considered a disorder of gut-brain-gut communication (Black et al., 2020a; Mayer et al., 2022).

INTEROCEPTIVE GUT-BRAIN COMMUNICATION: IMPACT ON MENTAL PROCESSES

The complex gut-brain-gut communication network contributes to interoception (Mayer, 2011; Holzer, 2017), a process that integrates information from all internal organs to impact on various domains of brain activity and behavior. In view of the neuronal and endocrine messaging modes of the signaling pathways it is obvious that interoceptive processes take place both at the conscious and subconscious level (Chen et al., 2021). Neuroanatomical and functional imaging studies have provided a detailed mapping of the brain regions in which interoceptive input is received, integrated and distributed to output relays. Whether delivered through neuronal or endocrine pathways, interoceptive information is first processed in subcortical structures of the CNS such as the spinal cord, brainstem and thalamus before it is passed on to higher brain regions including the hypothalamus, insula, anterior cingulate cortex, and somatosensory cortex (Critchley and Harrison, 2013; Wang et al., 2019; Chen et al., 2021). In this way interoceptive signals inform not only regulatory functions of the CNS to maintain internal homeostasis, but also influence feelings (mood, affect, emotion) and their valence as well as motivational and cognitive processes related to preferences, beliefs and decision-making (**Figure 1**) (Critchley and Harrison, 2013; Wang et al., 2019; Chen et al., 2021). In an experimental setting it has been shown that interoceptive signals from the gut carried by the vagus nerve support memory function in the hippocampus (Suarez et al., 2018).

Particular implications of gut-derived interoceptive signals in mental activity can be deduced from its disturbance in IBS which commonly is associated with visceral hyperalgesia as well as hypersensitivity to emotional challenge. A shown by functional imaging studies, painful rectal distension in healthy controls activates brain regions associated with visceral sensation and interoceptive processing (thalamus, anterior insula, anterior midcingulate cortex), emotional arousal (perigenual anterior cingulate gyrus) and attention and modulation of arousal (inferior parietal, lateral and medial prefrontal cortex) (Tillisch et al., 2011). In IBS patients, the activation of brain regions associated with visceral sensation, interoceptive processing and emotional arousal is significantly increased (Tillisch et al., 2011). In addition, psychological stress in IBS patients enhances the neuronal activation, which rectal distension induces in brain regions such as the insula, midcingulate cortex and ventrolateral prefrontal cortex, to a larger degree than in healthy controls (Elsenbruch et al., 2010). Likewise, IBS patients respond to contextual threat with increased neuronal activity within the salience, attention, default mode and emotional arousal networks within the CNS as compared with healthy controls, which appears to reflect the propensity of IBS subjects to overestimate the likelihood and severity of future abdominal pain (Hong et al., 2016). Furthermore, the hypersensitivity to rectal distension in IBS is related to changes in functional connectivity within resting-state networks associated with interoception, salience and sensory processing, changes that appear to be relevant to the hypervigilance and intestinal hyperalgesia seen in IBS patients (Icenhour et al., 2017). Meta-analyses have shown that patients with IBS present with significantly higher levels of anxiety and depression than healthy controls (Fond et al., 2014). Accordingly, cognitive behavioral therapy and gut-directed hypnotherapy have proved beneficial in a part of IBS patients (Peter et al., 2018; Black et al., 2020b), attesting to gut-brain-gut communication as a viable treatment target.

Background anxiety can strongly influence attitudes, beliefs and decisions, which is most evident in psychiatric disorders associated with generalized emotional disturbances. Decision-making depends on the computation of the value of available options, which in turn are a function of the environment and the internal state of the individual (Paulus and Yu, 2012). Engelmann et al. (2015) have shown that incidental anxiety disrupts the neural valuation of risky decision-making and shifts the valuation focus from possible positive consequences to anticipated negative consequences, a process in which the activity of the anterior insula plays a particular role. Transient anxiety states normally carry adaptive value since they may increase vigilance and attention to possible negative outcomes (Engelmann et al., 2015). This functional anxiety, however, can turn into a maladaptive

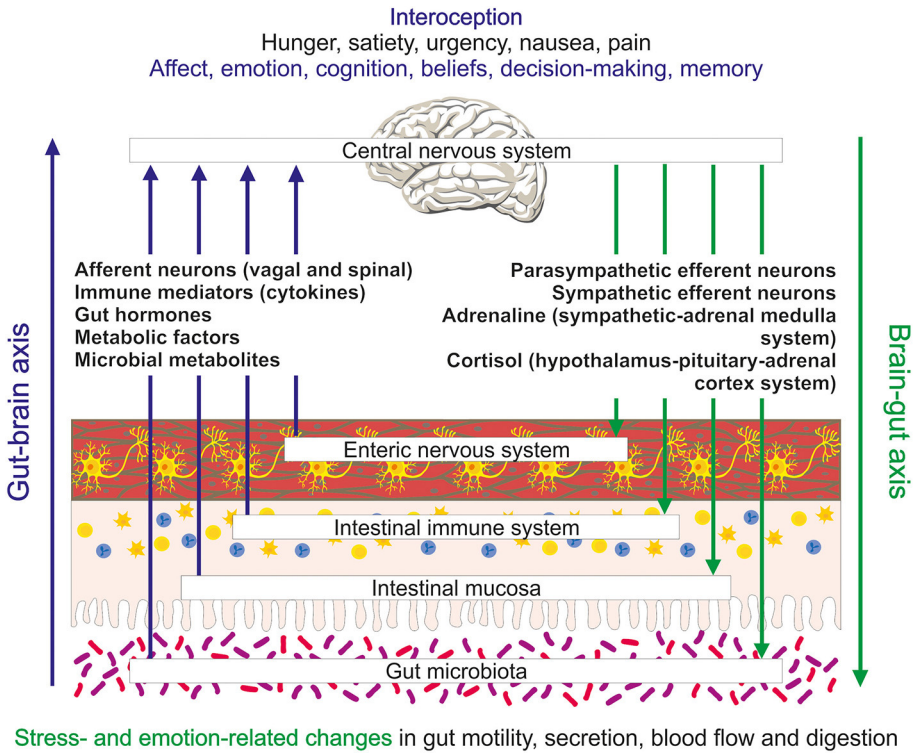


FIGURE 1 | Schematic illustration of the bidirectional exchange of information between the gut and brain and the interoceptive processing of this information in the brain. Several messenger systems including extrinsic afferent neurons, immune mediators, gut hormones, metabolic factors and microbial metabolites carry information from the gut to the brain. They elicit conscious sensations (e.g., hunger, satiety, urgency, nausea, and pain) and influence processes relevant to affect, emotion and cognition. Regulatory outputs from the brain to the gut include emotion- and stress-related changes in motility, secretion, blood flow and digestion.

state if anxious behavior is permanently adopted and becomes detached from the environment (Grupe and Nitschke, 2013). Affect can likewise have an adverse influence on decision-making (Paulus and Yu, 2012; Lerner et al., 2015). For instance, aversive affect appears to be a key source for irrational decision-making, especially with respect to trust in the context of social behavior (Engelmann et al., 2019).

INTEROCEPTIVE GUT SIGNALS AND "GUT FEELINGS": SCIENCE VERSUS BELIEF

The term "gut feelings" is a popular expression used in everyday language and refers to instinctive feelings, intuition, beliefs and decisions without rational underpinnings (Holzer, 2017). In this context, "gut feelings" are related to positive outcomes as exemplified by notions such as "gut feelings are guardian angels." The view that feelings originate in the gut may have also been fostered by the labeling of the enteric nervous system as "little brain" or "second brain" (Gershon, 1998). However, feelings and other mental capacities cannot be attributed to this nervous system in the gut, which is indispensable for the neural regulation of digestion (Holzer et al., 2001). Feelings or emotions do not originate in the gut but are generated in the brain, and the term "gut feelings" is a scientifically ill-defined and misleading expression. The impact of gut-derived interoceptive signals and sensations on mental health can be either positive or negative, the latter being aptly exemplified by the neuropsychiatric disturbances in IBS. There is no scientific evidence based on validated data that "gut feelings" have the power to direct judgements and decisions such that they have primarily a beneficial or happy payoff. To the contrary, instantaneous gut sensations known in neurogastroenterology, such as abdominal pain, flatulence, diarrhea-related urgency and nausea, are rather distressful. Notwithstanding these opposing views, the bidirectional communication network between gut and brain and the process of gastrointestinal interoception provide a neurobiological frame to explain emotions, beliefs, judgements and decisions under the influence of signals from the gut (Mayer et al., 2022).

The term "gut feeling" has also entered the scientific literature, which was fuelled not only by the elucidation of the gut-brain communication network but also driven by the entry of the gut microbiota as a factor of the gut-brain axis. Research in experimental models has provided a wealth of information on how the vast microbial community in the gut can participate in gut-brain signaling and interact with the neuronal, immune, endocrine and metabolic messengers of the gut-brain axis (Cryan et al., 2019; Farzi et al., 2019; Gershon and Margolis, 2021; Hassan et al., 2022). However, evidence for a direct impact of the gut microbiota on emotional-affective and cognitive behavior in humans lags behind, and microbiota-directed interventions with

proven efficacy in the management of neuropsychiatric disease are not yet available (Dinan and Cryan, 2019; Federici et al., 2020; Simpson et al., 2020; Le Morvan de Sequeira et al., 2022). While changes in the composition and diversity of the gut microbiome are associated with many neurological and psychiatric disorders (Simpson et al., 2021), causal relationships between particular aberrations of the gut microbiome and particular disorders of the human brain remain to be delineated. Despite the insufficient evidence, the hype in microbiome research is also mirrored in the popular press, the vast majority of articles (>90%) reporting health benefits associated with the gut microbiome without mentioning the limitations of such claims (Marcon et al., 2021). "Hope or hype" has become a common phrase in biomedical research areas in which a research boost raises expectations and beliefs in health benefits that await to be fulfilled.

Research hypes also carry the risk of deviating to questionable conceptions. One example relates to the purported mediator of the microbiome-gut-brain axis, 5-hydroxytryptamine (5-HT, serotonin), which both in the scientific and lay press is sometimes said to be an interface between gut microbiota and brain and to act as a "feel-good hormone". 5-HT synthesized in distinct brain neurons can in fact sustain good mood, and drugs (selective serotonin reuptake inhibitors) targeting the cerebral 5-HT system are efficacious in depression and certain anxiety disorders. However, more than 90% of the body's 5-HT is produced in the gut, primarily in enterochromaffin cells, but also in enteric neurons. Although the gut microbiota can indirectly modify the synthesis of 5-HT in gut and brain through regulating the availability of its precursor L-tryptophan (Gheorghe et al., 2019; Legan et al., 2022), intestinal 5-HT is unlikely to contribute to the "feel-good" action of cerebral 5-HT because it does not pass the blood-brain barrier. To the contrary, an excess of 5-HT in the gut can elicit nausea and emesis associated with chemo- and radiotherapy, facilitate intestinal inflammation, mediate diarrhea associated with bacterial infection, and contribute to IBS-related pain (Gershon, 2013; Legan et al., 2022).

CONCLUSION

The gut-brain-gut communication network is part of the interoceptive circuits which enable the brain to sense and interpret the physiological condition in the body and regulate its autonomic and mental activity accordingly. While this relationship has become an important research area in neuroscience, it also provides an example where solid science is at risk going uncritical and fostering unproven conceptions and expectations. It is here that credition research can find fruitful ground to analyze the working of science at the interface of "hype or hope" and to understand how interoceptive signals from the gut impact on mental activity to influence affect, emotion, beliefs, predictions and decisions. In its interdisciplinary approach, credition research is relevant to many areas in which belief processes shape religious, social, societal, economic, legal as well as scientific and medical conceptions and expectations. In analyzing these relationships, credition research bears considerable responsibility to unveil the misinterpretation of scientific data and the neglect of their validity status, which champion unproven notions and predictions. The placebo and nocebo effects represent a particular outcome of belief processes in which a complex set of input information convinces the patient that a certain choice of treatment is better or worse than the other although scientific evidence indicates that they are equivalent in their action. Importantly, placebo and nocebo effects are real, and they work either way, influencing brain activity as indicated by functional imaging studies and altering organ function in the periphery (Meissner, 2014; Bingel et al., 2022).

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

This paper was funded by Dr. Rüdiger Seitz, via the Volkswagen Foundation, Siemens Healthineers, and the Betz Foundation. Siemens Healthineers was not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

ACKNOWLEDGMENTS

The author thanks Judith C.J. Holzer-Geissler and Ulrike Holzer-Petsche for drawing **Figure 1**.

REFERENCES

Almy, T. P., and Tulin, M. (1947). Alterations in colonic function in man under stress: experimental production of changes simulating the irritable colon. *Gastroenterology* 8, 616–626.

Bingel, U., Wiech, K., Ritter, C., Wanigasekera, V., Mhuircheartaigh, R., Lee, M. C., et al. (2022). Hippocampus mediates nocebo impairment of opioid analgesia through changes in functional connectivity. *Eur J Neurosci in press*. doi: 10.1111/ejn.15687

Black, C. J., Drossman, D. A., Talley, N. J., Ruddy, J., and Ford, A. C. (2020a). Functional gastrointestinal disorders: advances in understanding and management. *Lancet* 396, 1664–1674. doi: 10.1016/S0140-6736(20)32115-2

Black, C. J., Thakur, E. R., Houghton, L. A., Quigley, E. M. M., Moayyedi, P., and Ford, A. C. (2020b). Efficacy of psychological therapies for irritable bowel syndrome: systematic review and network meta-analysis. *Gut* 69, 1441–1451. doi: 10.1136/gutjnl-2020-321191

Chen, W. G., Schloesser, D., Arensdorf, A. M., Simmons, J. M., Cui, C., Valentino, R., et al. (2021). https://pubmed.ncbi.nlm.nih.gov/33378655/The emerging science of interoception: sensing, integrating, interpreting, and regulating signals within the self. *Trends Neurosci.* 44, 3–16. doi: 10.1016/j.tins.2020.10.007

Craig, A. D. (2002). How do you feel? Interoception: the sense of the physiological condition of the body. *Nat. Rev. Neurosci.* 3, 655–666. doi: 10.1038/nrn894

Critchley, H. D., and Harrison, N. A. (2013). Visceral influences on brain and behavior. *Neuron* 77, 624–638. doi: 10.1016/j.neuron.2013.02.008

Cryan, J. F., O'Riordan, K. J., Cowan, C. S. M., Sandhu, K. V., Bastiaanssen, T. F. S., Boehme, M., et al. (2019). The microbiota-gut-brain axis. *Physiol. Rev.* 99, 1877–2013. doi: 10.1152/physrev.00018.2018

Dinan, T. G., and Cryan, J. F. (2019). Gut microbes and depression: still waiting for godot. *Brain Behav. Immun.* 79, 1–2. doi: 10.1016/j.bbi.2019.02.007

Elsenbruch, S., Rosenberger, C., Bingel, U., Forsting, M., Schedlowski, M., Gizewski, E. R. (2010). Patients with irritable bowel syndrome have altered emotional modulation of neural responses to visceral stimuli. *Gastroenterology* 139, 1310–1319. doi: 10.1053/j.gastro.2010.06.054

Engelmann, J. B., Meyer, F., Fehr, E., and Ruff, C. C. (2015). Anticipatory anxiety disrupts neural valuation during risky choice. *J. Neurosci.* 35, 3085–3099. doi: 10.1523/JNEUROSCI.2880-14.2015

Engelmann, J. B., Meyer, F., Ruff, C. C., and Fehr, E. (2019). The neural circuitry of affect-induced distortions of trust. *Sci. Adv.* 5, eaau3413. doi: 10.1126/sciadv.aau3413

Farzi, A., Hassan, A. M., Zenz, G., and Holzer, P. (2019). Diabesity and mood disorders: Multiple links through the microbiota-gut-brain axis. *Mol. Aspects Med.* 66, 80–93. doi: 10.1016/j.mam.2018.11.003

Federici, S., Suez, J., and Elinav, E. (2020). Our microbiome: On the challenges, promises, and hype. *Results Probl. Cell. Differ.* 69, 539–557. doi: 10.1007/978-3-030-51849-3_20

Fond, G., Loundou, A., Hamdani, N., Boukouaci, W., Dargel, A., Oliveira, J., et al. (2014). Anxiety and depression comorbidities in irritable bowel syndrome (IBS): A systematic review and meta-analysis. *Eur. Arch. Psychiatry Clin. Neurosci.* 264, 651–660. doi: 10.1007/s00406-014-0502-z

Gershon, M. D. (1998). *The Second Brain*. New York, NY: HarperCollins. ISBN-10:0060182520

Gershon, M. D. (2013). 5-Hydroxytryptamine (serotonin) in the gastrointestinal tract. *Curr. Opin. Endocrinol. Diabetes Obes.* 20, 14–21. doi: 10.1097/MED.0b013e32835bc703

Gershon, M. D., and Margolis, K. G. (2021). The gut, its microbiome, and the brain: Connections and communications. *J. Clin. Invest.* 131, e143768. doi: 10.1172/JCI143768

Gheorghe, C. E., Martin, J. A., Manriquez, F. V., Dinan, T. G., Cryan, J. F., and Clarke, G. (2019). Focus on the essentials: tryptophan metabolism and the microbiome-gut-brain axis. *Curr. Opin. Pharmacol.* 48: 137–145. doi: 10.1016/j.coph.2019.08.004

Grupe, D. W., and Nitschke, J. B. (2013). Uncertainty and anticipation in anxiety: an integrated neurobiological and psychological perspective. *Nat. Rev. Neurosci.* 14, 488–501. doi: 10.1038/nrn3524

Hassan, A. M., Mancano, G., Kashofer, K., Liebisch, G., Farzi, A., Zenz, G., et al. (2022). Anhedonia induced by high-fat diet in mice depends on gut microbiota and leptin. *Nutr. Neurosci.* 25, 299–312. doi: 10.1080/1028415X.2020.1775150

Holzer, P. (2017). "Interoception and gut feelings: unconscious body signals' impact on brain function, behavior and belief processes," in *Processes of Believing: The Acquisition, Maintenance, and Change in Creditions*, eds H. F. Angel, L. Oviedo, R. F. Paloutzian, A. L. C. Runehov, R. J. Seitz (Berlin: Springer), 435–442. doi: 10.1007/978-3-319-50924-2

Holzer, P., Hassan, A. M., Jain, P., Reichmann, F., and Farzi, A. (2015). Neuroimmune pharmacological approaches. *Curr. Opin. Pharmacol.* 25, 13–22. doi: 10.1016/j.coph.2015.09.003

Holzer, P., Schicho, R., Holzer-Petsche, U., and Lippe, I. T. (2001). The gut as a neurological organ. *Wien. Klin. Wochenschr.* 113, 647–660

Hong, J. Y., Naliboff, B., Labus, J. S., Gupta, A., Kilpatrick, L. A., Ashe-McNalley, C., et al. (2016). Altered brain responses in subjects with irritable bowel syndrome during cued and uncued pain expectation. *Neurogastroenterol. Motil.* 28, 127–138. doi: 10.1111/nmo.12710

Icenhour, A., Witt, S. T., Elsenbruch, S., Lowén, M., Engström, M., Tillisch, K., et al. (2017). Brain functional connectivity is associated with visceral sensitivity in women with irritable bowel syndrome. *Neuroimage Clin.* 15, 449–457. doi: 10.1016/j.nicl.2017.06.001

Kaelberer, M. M., Rupprecht, L. E., Liu, W. W., Weng, P., and Bohórquez, D. V. (2020). Neuropod cells: the emerging biology of gut-brain sensory transduction. *Annu. Rev. Neurosci.* 43, 337–353. doi: 10.1146/annurev-neuro-091619-022657

Le Morvan de Sequeira, C., Hengstberger, C., Enck, P., and Mack, I. (2022). Effect of probiotics on psychiatric symptoms and central nervous system functions in human health and disease: a systematic review and meta-analysis. *Nutrients* 14, 621. doi: 10.3390/nu14030621

Legan, T. B., Lavoie, B., and Mawe, G. M. (2022). Direct and indirect mechanisms by which the gut microbiota influence host serotonin systems. *Neurogastroenterol. Motil.* e14346. doi: 10.1111/nmo.14346

Lerner, J. S., Li, Y., Valdesolo, P., and Kassam, K. S. (2015). Emotion and decision making. *Annu. Rev. Psychol.* 66, 799–823. doi: 10.1146/annurev-psych-010213-115043

Marcon, A. R., Turvey, S., and Caulfield, T. (2021). 'Gut health' and the microbiome in the popular press: a content analysis. *BMJ Open* 11, e052446. doi: 10.1136/bmjopen-2021-052446

Mayer, E. A. (2000). The neurobiology of stress and gastrointestinal disease. *Gut* 47, 861–869. doi: 10.1136/gut.47.6.861

Mayer, E. A. (2011). Gut feelings: The emerging biology of gut-brain communication. *Nat. Rev. Neurosci.* 12, 453–466. doi: 10.1038/nrn3071

Mayer, E. A., Nance, K., and Chen, S. (2022). The gut-brain axis. *Annu. Rev. Med.* 73, 439–453. doi: 10.1146/annurev-med-042320-014032

McEwen, B. S. (2007). Physiology and neurobiology of stress and adaptation: central role of the brain. *Physiol. Rev.* 87, 873–904. doi: 10.1152/physrev.00041.2006

Meissner, K. (2014). Placebo responses on cardiovascular, gastrointestinal, and respiratory organ functions. *Handb. Exp. Pharmacol.* 225, 183–203. doi: 10.1007/978-3-662-44519-8_11

Paulus, M. P., and Yu, A. J. (2012). https://pubmed.ncbi.nlm.nih.gov/22898207/Emotion and decision-making: affect-driven belief systems in anxiety and depression. *Trends Cogn. Sci.* 16, 476–483. doi: 10.1016/j.tics.2012.07.009

Perez-Burgos, A., Mao, Y. K., Bienenstock, J., and Kunze, W. A. (2014). The gut-brain axis rewired: adding a functional vagal "sensory synapse". *FASEB J.* 28, 3064–3074. doi: 10.1096/fj.13-245282

Peter, J., Tran, U. S., Michalski, M., and Moser, G. (2018). The structure of resilience in irritable bowel syndrome and its improvement through hypnotherapy: cross-sectional and prospective longitudinal data. *PLoS One* 13, e0202538. doi: 10.1371/journal.pone.0202538

Seitz, R. J. (2022). Believing and beliefs – neurophysiological underpinnings. *Front. Behav. Neurosci.* 16, 880504. doi: 10.3389/fnbeh.2022.880504

Seitz, R. J., Angel, H. F., and Paloutzian, R. F. (2018). From believing to belief: a general theoretical model. *J. Cogn. Neurosci.* 30, 1254–1264. doi: 10.1162/jocn_a_01292

Shea-Donohue, T., Cook, I. J., de Giorgio, R., Tonini, M., Dent, J., Costa, M., et al. (2005). A teaching module on irritable bowel syndrome. *Neurogastroenterol. Motil.* 17(Suppl 3), 20–40. doi: 10.1111/j.1365-2982.2005.00713.x

Simpson, C. A., Diaz-Arteche, C., Eliby, D., Schwartz, O. S., Simmons, J. G., and Cowan, C. S. M. (2021). The gut microbiota in anxiety and depression - a systematic review. *Clin. Psychol. Rev.* 83, 101943. doi: 10.1016/j.cpr.2020.101943

Simpson, C. A., Schwartz, O. S., and Simmons, J. G. (2020). The human gut microbiota and depression: widely reviewed, yet poorly understood. *J. Affect. Disord.* 274, 73–75. doi: 10.1016/j.jad.2020.05.115

Suarez, A. N., Hsu, T. M., Liu, C. M., Noble, E. E., Cortella, A. M., Nakamoto, E. M., et al. (2018). Gut vagal sensory signaling regulates hippocampus function through multi-order pathways. *Nat. Commun.* 9, 2181. doi: 10.1038/s41467-018-04639-1

Taché, Y., Martinez, V., Million, M., and Wang, L. (2001). Stress and the gastrointestinal tract III. Stress-related alterations of gut motor function: role of brain corticotropin-releasing factor receptors. *Am. J. Physiol.* 280, G173–G177. doi: 10.1152/ajpgi.2001.280.2.G173

Tillisch, K., Mayer, E. A., and Labus, J. S. (2011). Quantitative meta-analysis identifies brain regions activated during rectal distension in irritable bowel syndrome. *Gastroenterology* 140, 91–100. doi: 10.1053/j.gastro.2010.07.053

Wang, X., Wu, Q., Egan, L., Gu, X., Liu, P., Gu, H., et al. (2019). Anterior insular cortex plays a critical role in interoceptive attention. *Elife* 8, e42265. doi: 10.7554/eLife.42265

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Holzer. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Psychology

OPINION
published: 27 July 2022
doi: 10.3389/fpsyg.2022.892829

Cognitive Bias: Phylogenesis or Ontogenesis?

Hans Van Eyghen [1,2]*

[1] Faculty of Religion and Theology, VU Amsterdam, Amsterdam, Netherlands, [2] Tilburg School of Catholic Theology, Tilburg University, Tilburg, Netherlands

Keywords: cognitive bias, evolutionary psychological science, agency detection, theory choice, bias

INTRODUCTION

The literature on cognitive biases is vastly expanding. The contribution of cognitive biases to the formation of beliefs and the process of believing (cf. Seitz et al., 2016; Angel, 2017;) is well documented. Well-documented examples are the confirmation bias (Mahoney, 1977), and the self-serving bias (Campbell and Sedikides, 1999). Most of the literature focuses on testing the existence and salience of various cognitive biases. Fewer authors focus on the *causes* of cognitive biases. This paper compares two mutually conflicting accounts of how cognitive biases arise. A first argues that (most) cognitive biases are part of the general human cognitive makeup, which is innate or emerges as humans mature. A second argues that cognitive biases are acquired throughout a human's lifespan and development. Below, I present examples of both accounts and reasons favoring each of both accounts.

A large number of definitions of "cognitive bias"[1] have been proposed. Some regard cognitive biases as *epistemic;* for example, as a "systematic pattern of deviation from norm or rationality in judgment" (Hasselton et al., 2005) or a "top-down, subjective directed perception" (Kahneman and Tversky, 1972). Others regard cognitive biases as forms of automatic cognition; for example as "automatic information processing." (Shiffrin and Schneider, 1977) or "Information processing without attention" (Payne and Gawronski, 2010). The definition used throughout this paper considers cognitive biases as skewed perceptions or skewed belief-formation. Because of cognitive biases, humans have a tendency for cognition to go in a particular direction, thereby giving rise to recurring patterns.

On most accounts of bias, humans usually remain unaware of their operations. While humans can be made aware through introspection or external information, most humans have a hard time explaining why their cognition is skewed in a particular direction.

In the next two sections I summarize two prominent accounts of how humans end up with cognitive biases. Caution should be made that the lines between both accounts are not clear-cut. Defenders of a phylogenistic account of cognitive bias usually agree that at least some biases are acquired throughout one's lifespan. The same holds for defenders of an ontogenistic account, although they allow for rather few innate cognitive biases. The difference is thus primarily one of focus, arguing that *most* or the *most salient* cognitive biases should be explained phylogenistically or ontogenistically.

BIAS IN PHILOGENESIS

A first account of the causes of cognitive biases argues that most cognitive biases result from the way the human mind is structured species-wide. Because of the way human minds and brains are, a number of biases arise. Some of these are innate, other gradually manifest as human brains mature[2].

[1] As the term designates, I will not take various biases for action into account in this paper.

[2] Some argue that phylogenetic enlargement of the brain could be related to the evolution of human beliefs (Seitz and Angel, 2020). The argument applies to the evolution of cognitive biases on a phylogenetic account as well.

OPEN ACCESS

Edited by:
Rüdiger J. Seitz, Heinrich Heine University of Düsseldorf, Germany

Reviewed by:
Hans Ferdinand Angel, University of Graz, Austria

***Correspondence:**
Hans Van Eyghen
hansvaneyghen@gmail.com

Specialty section:
This article was submitted to Theoretical and Philosophical Psychology, a section of the journal Frontiers in Psychology

Received: *09 March 2022*
Accepted: *13 June 2022*
Published: *27 July 2022*

Citation:
Van Eyghen H (2022) Cognitive Bias: Phylogenesis or Ontogenesis? Front. Psychol. 13:892829. doi: 10.3389/fpsyg.2022.892829

On this account, cognitive biases are ultimately encoded in the genome, which ultimately traces back to various evolutionary pressures.

While (some) biases require being triggered by sensory input on phylogenetic accounts, the response to that input is "preprogrammed" or predetermined by the human cognitive make-up. Sensory input is thus not the main determinant of why the bias originated.

One example of a bias that is commonly explained in this way is the bias to find food, heat and shelter (see Friston et al., 2012). Because most organisms require these very shortly after birth, the bias needs to be hard-wired or innate.

This line of reasoning has been applied to explain a cognitive bias for the detection of agents. Various authors have noted that humans are prone to see agency in natural or material phenomena. A famous example is the Heider-Simmel simulation (Heider and Simmel, 1944). The tendency to see agency even when such agency is absent under closer inspection would lie at the roots of animism or even religious beliefs according to some (e.g., Barrett, 2004).

Stewart Guthrie argues that our proneness to promiscuously perceive agency has its roots in evolutionary pressures. He argues that it was evolutionary beneficial for our distant ancestors to be highly sensitive for agency. Hypersensitivity compares favorably to lower sensitivity because low sensitivity increases the risk of missing one predator or rival human. Given the high threat, odds of doing so had better be diminished. Hypersensitivity diminishes the risk of missing predators or rivals greatly and only has minor costs (i.e., loss of time and energy) by comparison. For this reason, natural selection selected for a proneness to see agency on very limited evidence (Guthrie, 1993).

Guthrie does not discuss how evolutionary pressures led to a change in the human genome, which in turn leads to a cognitive bias. It is clear, however, that on Guthrie's account a bias for agency detection is not acquired throughout one's lifespan but part and parcel of the kind of mind or brain humans are endowed with. His account is therefore a clear example of a phylogenistic account of a cognitive bias.

BIAS IN ONTOGENESIS

Phylogenistic accounts of cognitive biases are arguably dominant in evolutionary psychology. Authors in cognitive neuroscience in particular tend to favor a different account where (most) cognitive biases are not hard-wired in the human brain but acquired because of the way the human mind engages with its environment. On such an ontogenetic account, humans are not born with the bias, not even in potentia.

Uncontroversial acquired biases are cultural biases like different responses to smiles or different levels of trust in various cultures (cf. Guiso et al., 2009). Skinner et al. (2020) argue that in-group biases are likely acquired by exposure to positive or negative responses to novel adults from out-groups. These biases are not ingrained within innate cognitive structures or the result of normal development thereof. Instead, they crucially depend on exposure to specific (sensory) input.

Ontogenetic accounts of the agency bias have been defended. Marc Andersen argues that the bias does not result from our evolved nature but depends on preexisting beliefs or priors that makes the presence of invisible agents more likely. Especially religious beliefs that invisible spirits or gods exist would raise the expectation of experiencing such beings in subjects. As a result, subjects with those beliefs would display a higher proneness to (over) detect agency (Andersen, 2017)[3].

Acquired biases need not be culturally specific. Some biases that recur cross-culturally can be the result of interactions with a similar environment by subjects with similar cognitive functions. Elsewhere I argued that a bias for agency (over) detection could result from common human processing of agents (Szocik and Van Eyghen, 2021). Given that human brains have limited computing power, brains have a hard time of keeping track of all features that indicate agency. Therefore, it is more efficient to focus on one or a few clear indicators of agency, like self-propelled movement and/or complex patterns. While such a simplification[4] allows human brains to quickly and efficiently detect agents, the flipside is increased proneness for misidentification. As Guthrie and other note, inanimate things occasionally appear to engage in self-propelled motion, like leaves rustling in the wind or branches falling from trees. A brain that focuses on self-propelled motion as an indicator of agency will therefore be biased to connect such movement to agency. Multiple cultural environments could give rise to the same simplification and therefore the bias could arise cross-culturally.

ONTOGENY OR PHYLOGENY?

Few, if any, authors are exclusivists with regard to a phylogenetic or ontogenetic genesis of biases. As noted, evolutionary psychologists lean toward accepting that more biases have a phylogenetic origin. Cognitive neuroscientists tend to accept a very limited number of biases of phylogenetic origin. As a result, there are conflicting accounts of a number of biases like the agency bias.

The existence of conflicting accounts suggests that both are underdetermined by the empirical data. On both accounts, biases have a similar phenomenology with recurring patterns in cognition that are hard to override. Nonetheless, both accounts predict some different empirical observations concerning biases. On an ontogenetic account, we would expect more variation depending on the (cultural) environment. An ontogenetic account also predicts more individual variation within groups.

Another observable difference is that phylogenetic biases are harder to override. Phylogenetic biases are regarded as a default state of the human cognitive system. While subjects can override this default state (for example, through rational deliberation or cognitive aides), the default state never disappears. When

[3]Andersen's account fits in a broader cognitive framework where the human mind is regarded as a self-organizing entity that builds an internal model of the world. The internal model holds statistical information on the likelyhood of encountering certain entities. The information can be updated when humans encounter mismatches between the inner model and sensory input (Friston, 2010).
[4]The simplification resembles feature reduction in statistical modeling in machine learning.

overriding factors lose their force, the default state will resurface. For example, some have argued that a bias to think of things teleologically or for a purpose resurfaces when subjects are put under time pressure (Kelemen et al., 2013) or forget about overriding information (Lombrozo et al., 2007). Phylogenetic biases thus repeatedly intrude or re-intrude on cognition. Given that ontogenetic biases are more malleable and display more variation, we would expect less intrusions of this kind if the bias were acquired[5].

Other evidence favoring a phylogenetic account would be evidence that a bias is present in very young children. Young children had little or almost no exposure to the sensory data needed for biases to take hold on an ontogenetic account. Therefore, evidence of a bias at a very young age is better explained as the result of the innate structures of their minds. A caveat must be made that sensory input already makes a substantial impact on children's minds from a very young age.

Contrary to what some suggest, evidence for a bias in non-human animals (e.g., Blanchard et al., 2014) does not necessarily favor a phylogenetic account. Like humans, some animals have learning-capacities that enable them to acquire biases through repeated exposure to similar sensory stimuli.

In the absence of evidence favoring either a phylogenetic or ontogenetic account, theoretical virtues can play a decisive role. An ontogenetic account is usually more parsimonious because it need not postulate anything beyond the plasticity of the human mind. An ontogenetic account also has more predictive power given that acquisitions of biases are easier to track than innate cognitive structures or evolutionary pressures.

[5]Ontogenetic biases could, however, also intrude on cognition if the biases get firmly entrenched. For a discussion, see: Szocik and Van Eyghen (2021).

CONCLUSION

I have surveyed two rivaling accounts of human cognitive bias. One puts its origins in the development of the human species and claims that the bias is part and parcel of natural human cognitive operations. The second states that the bias was acquired at some point in a human's development through specific sensory input and processing thereof. While I discussed a number of empirical traits that can help distinguish phylogenetic from ontogenetic biases, I argued that such evidence is often hard to come by. In the absence of such evidence, parsimony and predictive power generally favor an ontogenetic account.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

This paper was funded by the Volkswagen Foundation, Siemens Healthineers, and the Betz Foundation.

The authors declare that this study received funding from Siemens Healthineers. The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

REFERENCES

Andersen, M. (2017). Predictive coding in agency detection. *Religion Brain Behav.* 9, 65–84. doi: 10.1080/2153599X.2017.1387170

Angel, H. F. "Credition: From the question of belief to the question of believing", in *Processes of Believing: The Acquisition, Maintenance, and Change in Creditions.* Cham: Springer. (2017) p. 17–36. doi: 10.1007/978-3-319-50924-2_2

Barrett, J. L. (2004). *Why Would Anyone Believe in God? Cognitive Science of Religion Series*. Walnut Creek: Altamira Press.

Blanchard, T. C., Wilke, A., and Hayden, B. Y. (2014). "Hot-Hand Bias in Rhesus Monkeys." Journal of Experimental Psychology: Animal Learning and *Cognition* 40 (3):280. doi: 10.1037/xan0000033

Campbell, W. K., and Sedikides, C. (1999). Self-threat magnifies the self-serving-bias: a meta-analytic integration. *Rev Gen Psychol.* 3, 23–43. doi: 10.1037/1089-2680.3.1.23

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* 11, 127–138. doi: 10.1038/nrn2787

Friston, K., Thornton, C., and Clark, A. (2012). Free-energy minimization and the dark-room problem. *Front. Psycho.* 3, 130. doi: 10.3389/fpsyg.2012.00130

Guiso, L., Sapienza, P., and Zingales, L. (2009). Cultural biases in economic exchange? *Q. J. Econ.* 124, 1095–1131. doi: 10.1162/qjec.2009.124.3.1095

Guthrie, S. E. (1993). *Faces in the Clouds: A New Theory of Religion*. New York (N.Y.), Oxford: Oxford University Press.

Hasselton, M. G., Nettle, D., and Andrews, P. (2005). The evolution of cognitive bias. *Evol. Psychol.* 724–46. doi: 10.1002/9780470939376.ch25

Heider, F., and Simmel, M. (1944). An experimental study of apparent behavior. *Am. J. Psychol.* 57, 243—59. doi: 10.2307/1416950

Kahneman, D., and Tversky, A. (1972). subjective probability: a judgment of representativeness. *Cogn. Psychol.* 3, 430–454. doi: 10.1016/0010-0285(72)90016-3

Kelemen, D., Rottman, J., and Seston, R. (2013). Professional physical scientists display tenacious teleological tendencies: purpose-based reasoning as a cognitive default. *J. Exp. Psychol. Gen.* 142, 1074. doi: 10.1037/a0030399

Lombrozo, T., Kelemen, D., and Zaitchik, D. (2007). Inferring design evidence of a preference for teleological explanations in patients with Alzheimer's disease. *Adv. Cogn. Psychol.* 18, 999–1006. doi: 10.1111/j.1467-9280.2007.02015.x

Mahoney, M. J. (1977). Publication prejudices: an experimental study of confirmatory bias in the peer review system. *Cognit. Ther. Res.* 1, 161–175. doi: 10.1007/BF01173636

Payne, B. K., and Gawronski, B. (2010). "A history of implicit social cognition: where is it coming from? Where is it now? Where is it going", in *Handbook of Implicit Social Cognition: Measurement, Theory, and Applications* (New York, NY; London: Guilford Press), p. 1–15.

Seitz, R. J., and Angel, H. F. (2020). Belief formation–A driving force for brain evolution. *Brain Cognit.* 140, 105548. doi: 10.1016/j.bandc.2020.105548

Seitz, R. J., Paloutzian, R. F., and Angel, H. F. (2016). Processes of believing: where do they come from? What are they good for?. *F1000Research.* 5. doi: 10.12688/f1000research.9773.1

Shiffrin, R. M., and Schneider, W. (1977). Controlled and automatic human information processing: II. perceptual learning, automatic attending and a general theory. *Psychol. Rev.* 84, 127. doi: 10.1037/0033-295X.84.2.127

Skinner, A. L., Olson, K. R., and Meltzoff, A. N. (2020). Acquiring group bias: observing other people's nonverbal signals can create social group biases. *J. Personal. Social Psychol.* 119, 824. doi: 10.1037/pspi0000218

Szocik, K., and Van Eyghen, H. (2021). *Revising Cognitive and Evolutionary Science of Religion*. Religion as an Adaptation. London: Springer. doi: 10.1007/978-3-030-63516-9

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Van Eyghen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Pain Research

TYPE Opinion
PUBLISHED 29 July 2022
DOI 10.3389/fpain.2022.972169

Check for updates

OPEN ACCESS

EDITED BY
Rüdiger J. Seitz,
Heinrich Heine University of Düsseldorf, Germany

REVIEWED BY
Thorsten Kahnt,
National Institute on Drug Abuse Intramural Research Program, United States

*CORRESPONDENCE
Karin Meissner
karin.meissner@hs-coburg.de

SPECIALTY SECTION
This article was submitted to Pharmacological Treatment of Pain, a section of the journal Frontiers in Pain Research

RECEIVED 20 June 2022
ACCEPTED 11 July 2022
PUBLISHED 29 July 2022

CITATION
Meissner K (2022) Placebo, nocebo: Believing in the field of medicine. *Front. Pain Res.* 3:972169. doi: 10.3389/fpain.2022.972169

COPYRIGHT
© 2022 Meissner. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Placebo, nocebo: Believing in the field of medicine

Karin Meissner[1,2]*

[1]Division of Health Promotion, Coburg University of Applied Sciences and Arts, Coburg, Germany, [2]Medical Faculty, Institute of Medical Psychology, Ludwig-Maximilians-University Munich, Munich, Germany

KEYWORDS

placebo effect, nocebo effect, belief, expectation, hope, prediction error

Introduction

A medical treatment is regarded efficacious if it induces a larger improvement than an inert placebo treatment. The efficacy of the active treatment is usually tested in randomized placebo-controlled trials, which are expensive but necessary because also placebo treatments are associated with large improvements. This improvement is due in part to the occurrence of a "placebo effect."

The placebo effect is a genuine phenomenon that has been intensively researched in recent decades. A placebo treatment is by definition an inert treatment without specific ingredients, for example, a pill without pharmacologic ingredients. The placebo effect is best conceptualized as the effect of the informational context in which a (placebo or active) treatment is embedded and which consists of internal and external cues (1). External cues comprise, for example, the care provider's verbal suggestions about the effects of a treatment (e.g., "this drug is a powerful painkiller") as well as associated non-verbal cues conveyed through body language and facial expression. External cues include also the characteristics of a particular treatment, such as its invasiveness, price, color, and the medical setting in which it is applied (2). Internal cues that play a role in the formation of placebo effects comprise pre-existing expectancies, previous experience and the affective state of a patient. Internal and external cues interact with each other, and the resulting informational context can be considered the "active ingredient" of placebo interventions (1).

Typically, the informational treatment context translates into specific treatment expectations. Positive treatment expectations are thought to trigger placebo effects, that is, beneficial effects on health-related outcomes. When negative treatment expectations arise, so-called "nocebo effects" can occur, resulting, for example, in the occurrence or aggravation of symptoms. Treatment expectations are also able to modulate the effects of active treatments (3, 4). Negative treatment expectations are typically elicited by information about the risks of a treatment, communicated through healthcare professionals, medication leaflets, mass media, social media, and other patients (5). Nocebo effects have often been studied by analyzing the adverse events in the placebo arms in clinical trials. For example, a recent meta-analysis on the side effects of COVID-19 vaccination found that 76% of the systemic side effects after the first dose of vaccine, such as headache and fatigue, were also seen in the placebo groups, suggesting that the majority of systemic side effects were due to nocebo effects (6).

Placebo and nocebo effects can affect almost any medical symptom, including but not limited to pain, itch, nausea, depression, and motor symptoms (7). Notably, also physiological parameters, such as autonomic activity (8, 9) and plasma proteins (10) have been shown to be affected by placebo interventions. Furthermore, placebo effects can be surprisingly system specific: According to the content of the accompanying verbal suggestion, placebo interventions specifically affected gastric activity but not cardiovascular activity (11), and blood pressure but not gastric activity (12).

Neurobiological mechanisms

The neurobiological mechanisms underlying placebo effects differ depending on the conditions and paradigms used to induce placebo effects. According to their diversity, different neurochemical systems are known to be involved, including the opioid, dopamine, cholecystokinin, and oxytocin systems (1). For example, the opioid antagonist naloxone partially blocks placebo analgesia, whereas the cholecystokinin-antagonist proglumide inhibits the nocebo hyperalgesia, suggesting the involvement of opioidergic and cholecystokininergic pathways (13, 14). A recent meta-analysis of individual patient data from fMRI studies focusing on pain provided strong evidence for placebo-associated reductions of pain-related activity in areas linked to nociception and pain, such as the insular and thalamic regions. These changes, in turn, correlated with the magnitude of behavioral pain reduction (15). Effect sizes, however, were small, suggesting that further mechanisms underly placebo effects in pain. The meta-analysis also revealed increased activity in front-oparietal brain regions during placebo analgesia, particularly in the dorsolateral prefrontal cortex (DLPFC). This activation is thought to mirror the construction of top-down representations of context, including expectations and beliefs (15). The pivotal role of the DLPFC for placebo effects is nicely illustrated by an experimental study showing that the disruption of the DLPFC by repetitive transcranial magnetic stimulation completely blocked the placebo analgesic effect (16). In addition, the meta-analysis of placebo brain imaging studies showed a reduction of activity in brain areas related to negative affect (15). Accordingly, experimental evidence suggests that placebo effects on pain are partly mediated by reduced negative affect (17, 18), possibly induced by cognitive re-appraisal strategies (15). A further brain region frequently activated during placebo hypoanalgesia is the ventromedial prefrontal cortex (vmPFC), an area with a prominent role in decision making, valuation, and choice (1). VmPFC activation during placebo analgesia may reflect the occurrence of a decision bias evoked by the brain in an ambiguous situation. When positive treatment expectations, for example, let expect less pain while the nociceptive stimulus actually remains the same, a prediction error occurs. The brain may resolve this prediction error by a placebo hypoalgesic effect (19).

The majority of placebo effects are likely to be due to emotional re-appraisal strategies and cognitive-evaluative processes. Only very strong placebo interventions, such as induced by classical conditioning or powerful manipulations of belief, may affect early sensory processes in a significant manner (1).

Social neuropeptides and placebo effects

Allo-grooming in animals signals intense social relationships, and it has been postulated to constitute an important evolutionary trace of the placebo effect in humans (20–22). Indeed, empathetic behavior can enhance placebo effects. In a randomized controlled clinical trial on irritable bowel syndrome, for example, sham acupuncture was administered by a healthcare provider who was either instructed not to talk to the patients, or to interact with patients in an empathetic manner. Addition of empathy further enhanced the magnitude of the placebo effect induced by sham acupuncture (23). Furthermore, there is experimental evidence that neuropeptides released during social interactions, including oxytocin and vasopressin, can modulate placebo hypoalgesia (24, 25). For example, Colloca et al. (25) performed a randomized, placebo-controlled trial, in which nasal vasopressin agonists were administered to healthy volunteers before placebo analgesia was induced. The results showed that vasopressin remarkably enhanced the analgesic effect of the placebo intervention in women. By using plasma proteomics, we recently provided first evidence that the neuropeptides neurexin 1 (NRXN1), contactin-associated protein-like 4 (CNTNAP4), and reelin (RELN) play a role for the placebo effect in nausea (10). The cell adhesion molecules NRXN1 and CNTNAP4 are involved in mirror neuron activity and empathic behavior and have been linked to grooming behavior, and RELN is known to functionally interact with oxytocin. These preliminary results of an unbiased methodological approach (i.e., without a priori hypotheses) are promising, as they confirm previous findings that trust and a good doctor-patient relationship can improve medical outcomes and that such effects have a biological basis.

Open-label placebos

One of the most spectacular results of recent placebo research was the discovery that the open-label administration of placebos, where the patient is truthfully informed that the pill contains no pharmacological substance, produces a placebo effect. Since the first pilot study in patients with irritable bowel

syndrome (26), numerous trials have confirmed that open-label placebos can positively affect a variety of medical conditions, including but not limited to chronic low back pain, chronic knee pain, episodic migraine, allergic rhinitis, depression, attention deficit hyperactivity disorder, and cancer-related fatigue (27–32). There is even evidence that the beneficial effects of open-label placebos can last for several years (33).

The mechanisms underlying the effects of open-label placebos are largely unknown. A qualitative study in patients receiving open-label placebo within a clinical trial suggested that hope, rather than expectation, may play a role (19). While expectations refer to a relatively high (assumed) likelihood of the desired outcome and represent a rather cognitive construct, hope can also be present when the likelihood is very low and has often been conceptualized as an emotional state (34). Hope can drive patients to seek treatment even from a counterintuitive intervention such as open-label placebos (19).

Kaptchuk et al. (19) suggested that prediction error processing could explain the hypoalgesic effects of both deceptive and open-label placebos: In the case of deceptive placebo administration, positive expectations primarily lower the level of predicted pain, resulting in a prediction error which is resolved by the brain through a perceived hypoalgesic effect. In the case of OLP treatment, the placebo effect could be primarily due to reduced precision of the predicted pain signal, i.e., increased uncertainty resulting from the paradox information of receiving "substances that have no active ingredients." According to Bayesian brain models, the lowered precision of the "prior" (i.e., predicted pain) also leads to a prediction error, which in turn is resolved by a perceived hypoalgesic effect (19). Previous research on open-label placebos thus suggests that placebo effects can be elicited also in the absence of expectations, for example, when the patient is in an affective state of hope and increased uncertainty. Bayesian brain models provide a comprehensive model to explain both types of placebo effects.

The temporal dynamics of placebo effects

The multitude of mechanisms involved in placebo effects shows that this neurobiological phenomenon is complex and multifaceted. The temporal dynamics of placebo effects, however, have rarely been studied. Several authors suggested that perceived active treatment assignment may increase expectations, and thus placebo effects over time (35–37). In a randomized controlled trial in depression, for example, perceived treatment assignment affected symptom improvement only in the second half of the trial (37). "Active" placebo interventions that deliver non-specific sensory stimuli may be particularly useful in initiating such reinforcement processes. For example, adding electrotactile stimulation to a sham electrical nerve stimulation intervention for nausea significantly increased study participants' belief that they had received the "active" intervention. Although the magnitude of the placebo effect at the first placebo administration did not differ between the two placebo conditions, the difference in perceived treatment assignment could well lead to higher treatment expectations and thus placebo effects at subsequent placebo administrations (36). The long-lasting improvements in chronic low back pain observed during the 5-year follow-up of an open-label placebo study (33) furthermore suggests that placebo interventions can trigger strong and salient changes in patients' belief systems that may have long-term health effects. Altered cognitions, emotions and re-appraisal strategies, as well as changes in health behavior, may mediate such long-lasting placebo effects. Finally, also nocebo effects are most likely subject to changes over time, although empirical evidence in this area is limited due to ethical constraints.

Placebo effects and the process of believing

As outlined above, placebo research indicates that treatment expectations and related beliefs are not stable but are subject to change. Recently, a new area of research has emerged that aims to better understand beliefs as a function of "credition," that is, the "process of believing" (derived from the Latin verb "credere" - "to believe") (38). The process of believing is conceptualized as a basic brain function with neurophysiological underpinnings (39) that links past experience with predictions about the future and enables individuals to make sense of signals in the environment and ascribe personal meaning to them (38). Beliefs are the neural representations that result from the ongoing process of believing and can be reinforced and updated through learning processes. The model of credition thus shares many similarities with recent concepts in placebo research and offers a promising approach to better understand the dynamic formation of treatment-related beliefs and expectations as well as their clinical effects.

Concluding remarks

Placebo effects are not unique responses, but comprise a variety of mechanisms that differ between conditions and research paradigms. They rely on the brain's ability to actively integrate contextual information with prior experiences, conceptual knowledge, beliefs, and emotions, resulting in brain responses that promote health and well-being (1). There is considerable overlap with emerging concepts such as predictive coding and the process of believing. Integrating these concepts into placebo research could provide a better understanding of

the fluid nature of beliefs and expectations and their role in maintaining health and combating disease.

Author contributions

The author confirms being the sole contributor of this work and has approved it for publication.

Funding

This paper was funded by Rüdiger Seitz, *via* the Volkswagen Foundation, Siemens Healthineers, and the Betz Foundation. The funders were not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Wager TD, Atlas LY. The neuroscience of placebo effects: connecting context, learning and health. *Nat Rev Neurosci.* (2015) 16:403–18. doi: 10.1038/nrn3976

2. Meissner K, Linde K. Are blue pills better than green? How treatment features modulate placebo effects. *Int Rev Neurobiol.* (2018) 139:357–78. doi: 10.1016/bs.irn.2018.07.014

3. Colloca L, Lopiano L, Lanotte M, Benedetti F. Overt versus covert treatment for pain, anxiety, and Parkinson's disease. *Lancet Neurol.* (2004) 3:679–84. doi: 10.1016/S1474-4422(04)00908-1

4. Bingel U, Wanigasekera V, Wiech K, Mhuircheartaigh N, Lee R, Ploner MC, et al. The effect of treatment expectation on drug efficacy: imaging the analgesic benefit of the opioid remifentanil. *Sci Translat Med.* (2011) 3:70–4. doi: 10.1126/scitranslmed.3001244

5. Bingel U. Placebo 2.0: the impact of expectations on analgesic treatment outcome. *Pain.* (2020) 161(Suppl 1):S48–56. doi: 10.1097/j.pain.0000000000001981

6. Haas JW, Bender FL, Ballou S, Kelley JM, Wilhelm M, Miller FG, et al. Frequency of adverse events in the placebo arms of COVID-19 vaccine trials: A systematic review and meta-analysis. *JAMA Network Open.* (2022) 5:e2143955. doi: 10.1001/jamanetworkopen.2021.43955

7. Enck P, Bingel U, Schedlowski M, Rief W. The placebo response in medicine: minimize, maximize, or personalize? *Nat Rev Drug Discov.* (2013) 12:191–204. doi: 10.1038/nrd3923

8. Meissner K. The placebo effect and the autonomic nervous system: evidence for an intimate relationship. Philosophical Transactions of the Royal Society B. *Biol Sci.* (2011) 366:1808–17. doi: 10.1098/rstb.2010.0403

9. Meissner K. Placebo responses on cardiovascular, gastrointestinal, and respiratory organ functions. *Placebo.* (2014) 11:183–203. doi: 10.1007/978-3-662-44519-8_11

10. Meissner K, Lutter D, Von Toerne C, Haile A, Woods SC, Hoffmann V, et al. Molecular classification of the placebo effect in nausea. *PLoS ONE.* (2020) 15:e0238533. doi: 10.1371/journal.pone.0238533

11. Meissner K. Effects of placebo interventions on gastric motility and general autonomic activity. *J Psychosom Res.* (2009) 66:391–8. doi: 10.1016/j.jpsychores.2008.09.004

12. Meissner K, Ziep D. Organ-specificity of placebo effects on blood pressure. *Autonomic Neurosci.* (2011) 164:62–6. doi: 10.1016/j.autneu.2011.06.006

13. Colloca L, Benedetti F. Nocebo hyperalgesia: how anxiety is turned into pain. *Curr Opinion Anesthesiol.* (2007) 20:435–9. doi: 10.1097/ACO.0b013e3282b972fb

14. Meissner K, Bingel U, Colloca L, Wager TD, Watson A, Flaten MA. The placebo effect: advances from different methodological approaches. *J Neurosci.* (2011) 31:16117–24. doi: 10.1523/JNEUROSCI.4099-11.2011

15. Zunhammer M, Spisák T, Wager TD, Bingel U. Meta-analysis of neural systems underlying placebo analgesia from individual participant fMRI data. *Nature Commun.* (2021) 12:1–11. doi: 10.1038/s41467-021-21179-3

16. Krummenacher P. Prefrontal cortex mediated control of expectations in placebo analgesia. *Schmerz.* (2011) 25:440–3. doi: 10.1007/s00482-011-1049-9

17. Flaten MA, Aslaksen PM, Lyby PS, Bjørkedal E. The relation of emotions to placebo responses. *Philosophical Transact Royal Soc B Biol Sci.* (2011) 366:1818–27. doi: 10.1098/rstb.2010.0407

18. Geers AL, Faasse K, Guevarra DA, Clemens KS, Helfer SG, Colagiuri B. Affect and emotions in placebo and nocebo effects: what do we know so far? *Social Personal Psychol Compass.* (2021) 15:e12575. doi: 10.1111/spc3.12575

19. Kaptchuk TJ, Hemond CC, Miller FG. Placebos in chronic pain: evidence, theory, ethics, and use in clinical practice. *Bmj.* (2020) 370:1668. doi: 10.1136/bmj.m1668

20. Benedetti F. Placebo and the new physiology of the doctor-patient relationship. *Physiol Rev.* (2013) 93:1207–46. doi: 10.1152/physrev.00043.2012

21. Colloca L, Miller FG. How placebo responses are formed: a learning perspective. *Philosoph Transact Royal Soc B Biol Sci.* (2011) 366:1859–69. doi: 10.1098/rstb.2010.0398

22. Steinkopf L. The signaling theory of symptoms: an evolutionary explanation of the placebo effect. *Evol Psychol.* (2015) 13:1474704915600559. doi: 10.1177/1474704915600559

23. Kaptchuk TJ, Kelley JM, Conboy LA, Davis RB, Kerr CE, Jacobson EE, et al. Components of placebo effect: randomised controlled trial in patients with irritable bowel syndrome. *Bmj.* (2008) 336:999–1003. doi: 10.1136/bmj.39524.439618.25

24. Kessner S, Sprenger C, Wrobel N, Wiech K, Bingel U. Effect of oxytocin on placebo analgesia: a randomized study. *Jama.* (2013) 310:1733–5. doi: 10.1001/jama.2013.277446

25. Colloca L, Pine DS, Ernst M, Miller FG, Grillon C. Vasopressin boosts placebo analgesic effects in women: a randomized trial. *Biol Psych.* (2016) 79:794–802. doi: 10.1016/j.biopsych.2015.07.019

26. Kaptchuk TJ, Friedlander E, Kelley JM, Sanchez MN, Kokkotou E, Singer JP, et al. Placebos without deception: a randomized controlled trial in irritable bowel syndrome. *PLoS ONE.* (2010) 5:e15591. doi: 10.1371/journal.pone.0015591

27. Sandler AD, Bodfish JW. Open-label use of placebos in the treatment of ADHD: a pilot study. *Child Care Health Dev.* (2008) 34:104–10. doi: 10.1111/j.1365-2214.2007.00797.x

28. Kelley JM, Kaptchuk TJ, Cusin C, Lipkin S, Fava M. Open-label placebo for major depressive disorder: a pilot randomized controlled trial. *Psychother Psychosom.* (2012) 81:312–4. doi: 10.1159/000337053

29. Schaefer M, Harke R, Denke C. Open-label placebos improve symptoms in allergic rhinitis: a randomized controlled trial. *Psychother Psychosom*. (2016) 85:373–4. doi: 10.1159/000447242

30. Carvalho C, Caetano JM, Cunha L, Rebouta P, Kaptchuk TJ, Kirsch I. Open-label placebo treatment in chronic low back pain: a randomized controlled trial. *Pain*. (2016) 157:2766–72. doi: 10.1097/j.pain.0000000000000700

31. Olliges E, Stroppe S, Haile A, Reiß F, Malhis M, Funke SA, et al. Open-label placebo administration decreases pain in elderly patients with symptomatic knee osteoarthritis–A randomized controlled trial. *Front Psych*. (2022) 13:3497. doi: 10.3389/fpsyt.2022.853497

32. Kleine-Borgmann J, Schmidt K, Hellmann A, Bingel U. Effects of open-label placebo on pain, functional disability, and spine mobility in patients with chronic back pain: a randomized controlled trial. *Pain*. (2019) 160:2891–7. doi: 10.1097/j.pain.0000000000001683

33. Carvalho C, Pais M, Cunha L, Rebouta P, Kaptchuk TJ, Kirsch I. Open-label placebo for chronic low back pain: a 5-year follow-up. *Pain*. (2021) 162:1521–7. doi: 10.1097/j.pain.0000000000002162

34. Kube T, Rief W, Vivell MB, Schäfer NL, Vermillion T, Körfer K, et al. Deceptive and nondeceptive placebos to reduce pain: an experimental study in healthy individuals. *Clin J Pain*. (2020) 36:68–79. doi: 10.1097/AJP.0000000000000781

35. Meissner K, Distel H, Mitzdorf U. Evidence for placebo effects on physical but not on biochemical outcome parameters: a review of clinical trials. *BMC Med*. (2007) 5:1–11. doi: 10.1186/1741-7015-5-3

36. Aichner S, Haile A, Hoffmann V, Olliges E, Tschöp MH, Meissner K. The role of tactile stimulation for expectation, perceived treatment assignment and the placebo effect in an experimental nausea paradigm. *Front Neurosci*. (2019) 13:1212. doi: 10.3389/fnins.2019.01212

37. Laferton JA, Vijapura S, Baer L, Clain AJ, Cooper A, Papakostas G, et al. Mechanisms of perceived treatment assignment and subsequent expectancy effects in a double blind placebo controlled RCT of major depression. *Front*. (2018) 9:424. doi: 10.3389/fpsyt.2018.00424

38. Seitz RJ, Angel HF. Belief formation–A driving force for brain evolution. *Brain Cogn*. (2020) 140:105548. doi: 10.1016/j.bandc.2020.105548

39. Seitz RJ. Believing and beliefs-neurophysiological underpinnings. *Front Behav Neurosci*. (2022) 16:880504. doi: 10.3389/fnbeh.2022.880504

frontiers | Frontiers in Behavioral Neuroscience

OPINION
published: 18 July 2022
doi: 10.3389/fnbeh.2022.926742

Revealing the Cognitive Neuroscience of Belief

Michael H. Connors[1]* *and Peter W. Halligan*[2]

[1] Centre for Healthy Brain Ageing, University of New South Wales, Sydney, NSW, Australia, [2] School of Psychology, Cardiff University, Cardiff, United Kingdom

Keywords: belief, belief formation, cognition, cognitive neuropsychiatry, credition, delusion, psychosis

INTRODUCTION

Beliefs are convictions about what we accept as true. They provide the fundamental framework that we use to understand and engage meaningfully with the world. They also serve important social functions, such as in identity, relationships, and group coordination. Despite their personal and social significance, beliefs as psychological constructs have been largely neglected in empirical studies until recently. In previous work, we noted how studying delusion—defined as a pathological form of belief—provide a unique window to better understanding belief. Drawing on this approach and other psychological disciplines, we proposed a number of core functions and properties of belief. We also outlined a provisional five-stage cognitive model of belief formation. This paper provides an overview and discusses the implications of this account for psychology and cognitive neuroscience. In particular, we suggest that the five-stage model offers a tentative conceptual structure that could help foster future interdisciplinary research and render beliefs more tractable for scientific study.

OPEN ACCESS

Edited by:
Hans Ferdinand Angel, University of Graz, Austria

Reviewed by:
Yukiori Goto, Kyoto University, Japan

***Correspondence:**
Michael H. Connors
m.connors@unsw.edu.au

Specialty section:
This article was submitted to Emotion Regulation and Processing, a section of the journal Frontiers in Behavioral Neuroscience

Received: *23 April 2022*
Accepted: *24 June 2022*
Published: *18 July 2022*

Citation:
Connors MH and Halligan PW (2022) Revealing the Cognitive Neuroscience of Belief. Front. Behav. Neurosci. 16:926742. doi: 10.3389/fnbeh.2022.926742

STUDYING BELIEF

The notion of belief is frequently invoked and indeed assumed in everyday life and across all academic and clinical disciplines. Clarifying the construct is critical, in particular, for cognitive psychology given its responsibility to characterize the mental processes underpinning how we, as self-embodied individuals, understand and engage with others and our physical environment (Connors and Halligan, 2015). Despite this, empirical research and theoretical discussions within psychology remain limited. This has likely been driven by difficulties operationalizing this ubiquitous and nebulous term. Philosophical debates about the nature of belief continue (Schwitzgebel, 2010); folk conceptions vary (Pechey and Halligan, 2012b); and cognitive accounts have not been available until recently. Such issues, however, can be overcome. Philosophical debates and folk conceptions need not preclude empirical study (Bell et al., 2006) and recent theoretical developments offer greater clarity around research directions (Connors and Halligan, 2020).

A related challenge for research has been the inherent complexity of belief. Beliefs exist within broader networks of related beliefs (Quine and Ullian, 1970), making discrete beliefs difficult to study in isolation. Beliefs also interact with many lower-level cognitive processes, such as attention, perception, and memory. Given such close inter-relationships with automatic cognitive processes, Fodor (1983) argued that belief could not be decomposed into discrete independent subcomponents (modules) or localized neuroanatomically, limiting the viability of scientific study. These pessimistic accounts have been challenged over recent decades. While some suggest that it is premature to accept that belief is non-modular (Coltheart, 2017), others note that, even if this turns out to be the case, it need not follow that scientific study is impossible (Murphy, 2019). Scientific methods have been effectively applied to many complex systems and other forms of higher-order cognition, suggesting that analogous methods could be developed for the unique subject matter (Connors and Halligan, 2020).

One method that offers promise when addressing these practical challenges is the study of delusions. Delusions offer salient examples of pathological belief and often reflect relatively circumscribed dysfunction within an individual's cognitive system. By careful study, one can identify specific contributory factors and their impact. Such study, in turn, can offer insights into the cognitive processes involved in other delusions and belief formation more generally (Connors and Halligan, 2015, 2017, 2020). This approach is known as cognitive neuropsychiatry—a discipline that seeks to explain neuropsychiatric symptoms in terms of disruptions or damage to normal cognitive processes (Halligan and David, 2001).

The approach can be briefly illustrated when applied to the Capgras delusion—a false belief that a familiar person has been replaced by an impostor. Research has found that multiple patients with this delusions have a deficit in their autonomic response to familiar faces (Ellis et al., 1997). This deficit could plausibly lead to an unexpected sense of unfamiliarity around others and hence a conclusion that a familiar person is an impostor (Ellis and Young, 1990). Similar accounts can be offered for the content of other delusions, reflecting the more general point that delusions may arise from an attempt to explain unusual sensory data. Some patients, however, experience anomalous data without developing a delusion. This suggests the need for another factor—such as a deficit in belief evaluation (Coltheart et al., 2011)—to explain why some patients accept the delusion and others do not. Whilst aspects of this account remain subject to discussion, the example highlights the broader potential to examine cognitive processes contributing to delusional belief.

A FIVE-STAGE ACCOUNT

Based on evidence from delusions and other psychological disciplines, we previously identified several core functions of beliefs. These include providing a consistent representation of our social and physical world; offering an explanatory framework; coordinating lower-level cognitive processes; and facilitating social functions, such as identity, relationships, and group coordination (Connors and Halligan, 2015). We also identified a range of dimensions of belief, such as their origins, conviction, stability, conscious awareness, and impact.

Against this background, we outlined a tentative five-stage cognitive model of belief formation (Connors and Halligan, 2015, 2017, 2020). This noted that beliefs are likely to arise in response to a precursor, a distal trigger for the belief's content (**Figure 1**). Between the precursor and the belief, at least two intermediate stages are needed: firstly, ascribing meaning to the precursor and, secondly, evaluating the proposed meaning in terms of whether it meets criteria for belief. After a belief is formed, a fifth stage is the effect the belief then has on subjective experience and other cognitive processes, including other beliefs. This overall account is not committed to modularity and individual stages are likely underpinned by a wide range of automatic and unconscious cognitive processes (Oakley and Halligan, 2017). It is nevertheless possible to characterize these broad stages in more detail.

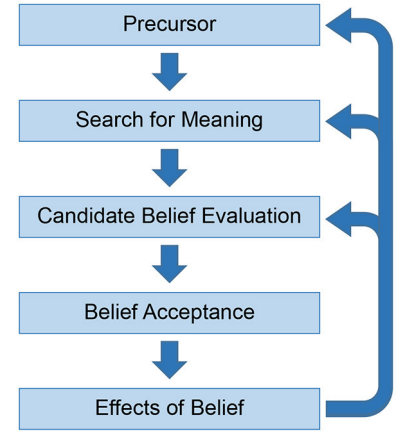


FIGURE 1 | Five stages of belief formation.

The first stage is a precursor that operates as a distal trigger for a belief's content. This could involve sensory input, particularly if unexpected or otherwise salient. It could, however, also take other forms, such as communication from trusted others. Indeed, many beliefs, including pathological forms, appear to arise from accepting social communicated ideas, rather than direct sensory experience (Sperber, 2009). A further form of precursor may be introspection on memories, imagery, or pre-existing beliefs, which can likewise occur without immediate sensory input. For delusions, source monitoring errors—failures to identify the origins of internally-generated thoughts, memories, and actions—may provide an important source of content (Johnson, 1988; Griffin and Fletcher, 2017).

The second stage is a search for meaning to interpret and explain the precursor. This draws heavily upon pre-existing beliefs and other relevant contextual information. As such, the resulting proto-beliefs can be highly personal and idiosyncratic. Interpretation is likely, in particular, to seek to preserve pre-existing beliefs for internal consistency and avoid dissonance. Interpretation is also likely to reflect particular attributional styles (habitual tendencies to explain events in certain ways); heuristics to reduce cognitive effort; emotion and mood; and social motivations (e.g., preserving a positive sense of self and maintaining relationships and group ties). Such processes shape the content of beliefs beyond what is specified by the precursor itself.

The third stage involves evaluating the proto-beliefs. This is likely based on at least two key criteria, namely observational adequacy (the degree to which the belief explains the precursor) and doxastic conservatism (consistency with pre-existing beliefs; Stone and Young, 1997; Mckay, 2012). The latter tendency is particularly important for maintaining internal consistency, so conflicting accounts are likely subjected to intense scrutiny. For delusions, disruptions in belief evaluation may give rise to implausible content by allowing hypotheses to be accepted without adequate examination. Such disruptions, however, are not necessary for all delusions. Supportive pre-existing beliefs

and/or cognitive biases within the subject's own community could also allow many unusual beliefs to be accepted. Indeed, once delusions are formed, belief evaluation may serve to reject alternative, non-delusional accounts to maintain internal consistency.

The fourth stage is activation of the new belief. This will usually need to be co-located within a network of inter-related beliefs to be maintained. As already noted, beliefs vary in specific properties and multiple factors are likely to influence each of these. Of particular significance are a belief's conviction and influence on action. These features are likely to depend on similar criteria as those in belief evaluation, namely the belief's adequacy at predicting ongoing experience and congruence with other pre-existing beliefs. Both criteria, as well as a belief's salience, can vary to some degree across time and context, so it is possible that a belief's conviction and influence may similarly vary (Connors and Coltheart, 2011). Most beliefs, however, are likely to fit within a network of consistent and mutually supportive beliefs (Pechey and Halligan, 2012a; Seitz et al., 2018), so are likely to remain relatively stable.

The final stage is the impact the belief has on lower-level cognitive processes and broader subjective experience. In everyday life, beliefs are experienced as lived and typically not subject to decomposition, questioning, or reflection at the time. As representations of one's phenomenal world, beliefs strongly influence attributions and the deployment of lower-order processes, such as attention, perception, and memory, in a top-down manner. Whilst constrained by sensory data, beliefs bias cognitive processing, particularly when data are ambiguous, to align with the beliefs' predictions. Specific mechanisms remain contested, including the extent to which beliefs affect basic perception (Vetter and Newen, 2014). Nevertheless, the overall impact of beliefs on attributions and subjective experience is evident across many experimental paradigms (Hastorf and Cantril, 1954; Jones and Russell, 1980; Gilovich, 1991; Gregory, 1997; Irwin, 2009; Connors et al., 2015). As such, beliefs, including delusional forms, provide an incredibly powerful lens that shapes our experience, affecting what we attend to, perceive, remember, and consider plausible as an explanation. This, in turn provides further support for the belief and lead to the elaboration of related beliefs and broader world views.

IMPLICATIONS

Our account of belief formation is admittedly preliminary and underspecified. We consider it, however, to be parsimonious and helpful when trying to explain the heterogeneity of belief, including delusions and other anomalous forms. We also believe that it has sufficient detail to guide future research. Our five-stage account highlights, in particular, how belief formation can be functionally decomposed, independent of assumptions around cognitive architecture and modularity. This has relevance to other areas of psychology, cognitive neuroscience, and many other academic disciplines. It also provides a more comprehensive account of delusions than existing cognitive accounts, which have a number of significant empirical and theoretical limitations (Connors and Halligan, 2020, 2021a,b).

Research methodologies from many disciplines are relevant when elucidating the cognitive processes implicated. While studying delusions is likely to remain important, observational research of beliefs in the non-clinical population will be needed to better define characteristics of normality and dysfunction. Strongly-held beliefs with anomalous content—such as conspiracy theories and certain religious and political beliefs—may be particularly relevant in this respect and provide insight into developmental antecedents, personality factors, neuropsychological correlates, and social dynamics (Pechey and Halligan, 2011; Douglas et al., 2017). Experimental methods that directly alter belief, including associative learning, hypnosis (Oakley and Halligan, 2013; Connors, 2015), and social influence (Cialdini, 2021), are also likely to be important in clarifying psychological mechanisms.

A final challenge involves marrying the cognitive processes of belief to the underlying neurobiology (Bell and Halligan, 2013). Recent accounts have highlighted potential neurophysiological processes involved in believing ("credition"; Seitz et al., 2018). Importantly, however, neuroimaging and other investigative techniques depend in large part on the cognitive models and behavioral tasks used (Poldrack and Farah, 2015). As such, the five-stage account provides an initial cognitive framework to guide investigation. Our account also highlights challenges establishing specificity of associations given beliefs' heterogeneous properties; frequent coalescence around shared themes; and close connections with lower-level automatic cognitive systems. Contrary to Fodor, these challenges are not necessarily insurmountable, though care will be need to be taken in experimental designs and likely require converging methodologies. Computational modeling and predictive data-driven approaches may assist, though both similarly remain limited to some extent by the overarching cognitive framework used (Poldrack and Yarkoni, 2016). Progress in cognitive neuroscience is therefore likely to remain closely linked to elucidating belief's cognitive basis. Further clarification of both promises to offer important insights into consciousness, social processes, and ourselves.

AUTHOR CONTRIBUTIONS

MC and PH were involved in conceptualizing, writing, and revising this paper for critically meaningful content. Both authors approved the final version.

FUNDING

Publication fees were paid by Professor Rüdiger Seitz via the Volkswagen Foundation, Siemens Healthineers, and the Betz Foundation.

ACKNOWLEDGMENTS

We are grateful to Professor Rüdiger Seitz for inviting this paper and paying the publication fees.

REFERENCES

Bell, V., and Halligan, P. W. (2013). "The neural basis of abnormal personal belief," in *The Neural Basis of Human Belief Systems,* eds F. Krueger and J. Grafman (Hove, UK: Psychology Press), 191–224.

Bell, V., Halligan, P. W., and Ellis, H. D. (2006). "A cognitive neuroscience of belief," in *The Power of Belief: Psychosocial Influence on Illness, Disability and Medicine,* eds P. W. Halligan and M. Aylward (Oxford, UK: Oxford University Press), 3–20.

Cialdini, R. B. (2021). *Influence: The Psychology of Persuasion.* New York, NY: HarperCollins.

Coltheart, M. (2017). The assumptions of cognitive neuropsychology: reflections on Caramazza (1984, 1986). *Cogn. Neuropsychol.* 34, 397–402. doi: 10.1080/02643294.2017.1324950

Coltheart, M., Langdon, R., and Mckay, R. (2011). Delusional belief. *Annu. Rev. Psychol.* 62, 271–298. doi: 10.1146/annurev.psych.121208.131622

Connors, M. H. (2015). Hypnosis and belief: a review of hypnotic delusions. *Conscious. Cogn.* 36, 27–43. doi: 10.1016/j.concog.2015.05.015

Connors, M. H., Barnier, A. J., Langdon, R., and Coltheart, M. (2015). Hypnotic models of mirrored-self misidentification delusion: a review and an evaluation. *Psychol. Consciousness* 2, 430–451. doi: 10.1037/cns0000059

Connors, M. H., and Coltheart, M. (2011). On the behaviour of senile dementia patients vis-à-vis the mirror: Ajuriaguerra, Strejilevitch and Tissot (1963). *Neuropsychologia* 49, 1679–1692. doi: 10.1016/j.neuropsychologia.2011.02.041

Connors, M. H., and Halligan, P. W. (2015). A cognitive account of belief: a tentative roadmap. *Front. Psychol.* 5, 1588. doi: 10.3389/fpsyg.2014.01588

Connors, M. H., and Halligan, P. W. (2017). "Belief and belief formation: Insights from delusions," in *Processes of Believing: The Acquisition, Maintenance, and Change in Creditions*, eds H.-F. Angel, L. Oviedo, R. F. Paloutzian, A. L. C. Runehov, and R. J. Seitz (Cham: Springer International Publishing), 153–165.

Connors, M. H., and Halligan, P. W. (2020). Delusions and theories of belief. *Conscious. Cogn.* 81, 102935. doi: 10.1016/j.concog.2020.102935

Connors, M. H., and Halligan, P. W. (2021a). Delusions and disorders of self-experience. *Lancet Psychiatry* 8, 740–741. doi: 10.1016/S2215-0366(21)00244-3

Connors, M. H., and Halligan, P. W. (2021b). Phenomenology, delusions, and belief. *Lancet Psychiatry* 8, 272–273. doi: 10.1016/S2215-0366(21)00027-4

Douglas, K. M., Sutton, R. M., and Cichocka, A. (2017). The psychology of conspiracy theories. *Curr. Dir. Psychol. Sci.* 26, 538–542. doi: 10.1177/0963721417718261

Ellis, H. D., and Young, A. W. (1990). Accounting for delusional misidentifications. *Br. J. Psychiatry* 157, 239–248. doi: 10.1192/bjp.157.2.239

Ellis, H. D., Young, A. W., Quayle, A. H., and De Pauw, K. W. (1997). Reduced autonomic responses to faces in Capgras delusion. *Proc. R. Soc. B Biol. Sci.* 264, 1085–1092. doi: 10.1098/rspb.1997.0150

Fodor, J. A. (1983). *The Modularity of Mind: An Essay on Faculty Psychology*. Cambridge, MA: The MIT Press.

Gilovich, T. (1991). *How we Know What isn't so: The Fallibility of Human Reason in Everyday Life*. New York, NY: The Free Press.

Gregory, R. L. (1997). *Eye and Brain: The Psychology of Seeing.* Princeton, NJ: Princeton University Press.

Griffin, J. D., and Fletcher, P. C. (2017). Predictive processing, source monitoring, and psychosis. *Annu. Rev. Clin. Psychol.* 13, 265–289. doi: 10.1146/annurev-clinpsy-032816-045145

Halligan, P. W., and David, A. S. (2001). Cognitive neuropsychiatry: towards a scientific psychopathology. *Nat. Rev. Neurosci.* 2, 209–215. doi: 10.1038/35058586

Hastorf, A. H., and Cantril, H. (1954). They saw a game: a case study. *J. Abnorm. Soc. Psychol.* 49, 129–134. doi: 10.1037/h0057880

Irwin, H. J. (2009). *The Psychology of Paranormal Belief: A Researcher's Handbook.* Hertfordshire: University of Hertfordshire Press.

Johnson, M. K. (1988). "Discriminating the origin of information," in *Delusional Beliefs,* eds T. F. Oltmanns and B. A. Maher (New York, NY: John Wiley and Sons), 34–65.

Jones, W. H., and Russell, D. (1980). The selective processing of belief disconfirming information. *Eur. J. Soc. Psychol.* 10, 309–312. doi: 10.1002/ejsp.2420100309

Mckay, R. (2012). Delusional inference. *Mind Lang.* 27, 330–355. doi: 10.1111/j.1468-0017.2012.01447.x

Murphy, G. L. (2019). On Fodor's first law of the nonexistence of cognitive science. *Cogn. Sci.* 43, e12735. doi: 10.1111/cogs.12735

Oakley, D. A., and Halligan, P. W. (2013). Hypnotic suggestion: opportunities for cognitive neuroscience. *Nat. Rev. Neurosci.* 14, 565–576. doi: 10.1038/nrn3538

Oakley, D. A., and Halligan, P. W. (2017). Chasing the rainbow: the non-conscious nature of being. *Front. Psychol.* 8, 1924. doi: 10.3389/fpsyg.2017.01924

Pechey, R., and Halligan, P. (2011). The prevalence of delusion-like beliefs relative to sociocultural beliefs in the general population. *Psychopathology* 44, 106–115. doi: 10.1159/000319788

Pechey, R., and Halligan, P. (2012a). Using co-occurrence to evaluate belief coherence in a large non clinical sample. *PLoS ONE* 7, e48446. doi: 10.1371/journal.pone.0048446

Pechey, R., and Halligan, P. W. (2012b). Exploring the folk understanding of belief: identifying key dimensions endorsed in the general population. *J. Cogn. Cult.* 12, 81–99. doi: 10.1163/156853712X633947

Poldrack, R. A., and Farah, M. J. (2015). Progress and challenges in probing the human brain. *Nature* 526, 371–379. doi: 10.1038/nature15692

Poldrack, R. A., and Yarkoni, T. (2016). From brain maps to cognitive ontologies: informatics and the search for mental structure. *Annu. Rev. Psychol.* 67, 587–612. doi: 10.1146/annurev-psych-122414-033729

Quine, W. V., and Ullian, J. S. (1970). *The Web of Belief.* New York, NY: Random House.

Schwitzgebel, E. (2010). "Belief," in *The Stanford Encyclopedia of Philosophy,* ed E.N. Zalta (Stanford, CA: Stanford University).

Seitz, R.J., Paloutzian, R.F., and Angel, H.-F. (2018). From believing to belief: A general theoretical model. *J. Cogn. Neurosci.* 30, 1254–1264. doi: 10.1162/jocn_a_01292

Sperber, D. (2009). Culturally transmitted misbeliefs. *Behav. Brain Sci.* 32, 534–535. doi: 10.1017/S0140525X09991348

Stone, T., and Young, A. W. (1997). Delusions and brain injury: the philosophy and psychology of belief. *Mind Lang.* 12, 327–364. doi: 10.1111/1468-0017.00051

Vetter, P., and Newen, A. (2014). Varieties of cognitive penetration in visual perception. *Conscious. Cogn.* 27, 62–75. doi: 10.1016/j.concog.2014.04.007

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Connors and Halligan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Psychology

OPINION
published: 03 June 2022
doi: 10.3389/fpsyg.2022.929143

Control Over Believing and Doxastic Responsibility

*Aku Visala**

Department of Systematic Theology, Faculty of Theology, University of Helsinki, Helsinki, Finland

Keywords: responsibility, belief, control, moral responsibility, doxastic responsibility

The theory of credition suggests that we should use active and agential language about beliefs. Instead of "beliefs" we should talk about "believing". The approach encourages us to see believing as a "mental activity" or "dynamic activity" (Angel et al., 2017). These expressions raise an important question: is "believing" something that human agents do or does it happen to us? Or to put the question in another way: are our beliefs results of our own agency or not? In what follows, I will examine this question from the point of view of responsibility.

DOXASTIC RESPONSIBILITY AND CONTROL

For our responsibility attributions, it matters greatly if our beliefs can be said to be products of our agency. Consider the following argument:

(1) If epistemic responsibility attributions (e.g., blame and praise) are appropriate, we have voluntary control over our believing.
(2) We have no voluntary control over our believing.
(3) Therefore, epistemic responsibility attributions are not appropriate.

Premises (1) and (2) seem quite plausible. We humans assume that in order for responsibility attributions to be appropriate the target of those attributions must be under the control of the agent. We think it unjust to blame a person for an action that she did not control.

Furthermore, we often take beliefs to be analogous to actions; we treat beliefs as an expression of an agent's agency. If someone holds an irrational belief, we blame the person for failing to achieve expected epistemic standards. Notice, that this attribution of blame assumes that whether a person fulfills or fails those epistemic standards falls under the control of the person. Again, if the person exercised no control over her epistemic life, there would be no point in blaming her for the failure.

The second premise seems plausible as well. Choosing one's beliefs seems, after all, impossible. I cannot decide or choose to believe whether there is a computer is in front of me. If I see a computer before me, I believe it. If I do not see it, I do not believe it. We seem to be passive recipients of beliefs rather than authors of them. Our cognitive system produces beliefs without our conscious input on the basis of how it perceives the world. Since we do not choose our beliefs, we cannot be blamed or praised for them either.

Facing this dilemma, we have three options:

(a) Doxastic voluntarism.
(b) Doxastic involuntarism + ground epistemic responsibility judgments on something else than control over believing.
(c) Doxastic involuntarism + reject epistemic responsibility and revise our practices accordingly.

OPEN ACCESS

Edited by:
Rüdiger J. Seitz, Heinrich Heine University of Düsseldorf, Germany

Reviewed by:
Tim Kraft, University of Regensburg, Germany

***Correspondence:**
Aku Visala
aku.visala@helsinki.fi

Specialty section:
This article was submitted to Theoretical and Philosophical Psychology, a section of the journal Frontiers in Psychology

Received: *26 April 2022*
Accepted: *19 May 2022*
Published: *03 June 2022*

Citation:
Visala A (2022) Control Over Believing and Doxastic Responsibility. Front. Psychol. 13:929143. doi: 10.3389/fpsyg.2022.929143

Most philosophers tend to gravitate toward options (b) and (c)[1]. Against this, I want to defend the plausibility of option (a), doxastic voluntarism. I will suggest that because philosophers have had such a high standard for what voluntary control requires, they have mistook believing as a passive process that does not involve the person's agency at all. We might not choose or decide to acquire most of our beliefs, but that does not mean believing is a passive process outside of our control.

TWO ARGUMENTS FOR DOXASTIC INVOLUNTARISM

Consider one practical and empirical argument against doxastic voluntarism. Robert Audi has argued that an evolved creature would be highly unlikely to develop a cognitive system that could acquire and maintain beliefs at will (Audi, 2015, p. 34–42). This is because holding a distinction between cognitive systems that represent the world and cognitive systems that facilitate and maintain the organism's aims and goals is crucial. If the organism fails to distinguish between how the world actually is and how the organism wants it to be, it will never achieve its goals. As a consequence, the processes of the "intellect" (getting at true beliefs) and the "will" (practical reasoning) will become independent over time.

More conceptual argument against doxastic voluntarism comes from Bernard Williams[2]. For Williams, beliefs are intrinsically aimed at truth. As such, they must be caused by something that is truth tracking or truth-apt, namely, something that makes a belief true (or false). This can be a piece of evidence, like perception, inference from knowledge, memory or something else. Consider now the possibility that I acquire a belief simply by forming an intention to acquire such a belief. If I know that I have acquired a belief simply because I have formed the intention to do so, I also know that this particular belief was not caused by a truth-tracking reason. Intentions to acquire a belief are not truth-tracking. So, if I know that I have decided to adopt a belief, I also know that the belief in question is not a product of a truth-tracking reason. This makes believing at will incoherent.

GUIDANCE CONTROL

The two previous arguments strike true to me. It seems that synchronically deciding to believe something is psychologically very difficult and conceptually impossible. It does not follow from this, however, that we cannot exercise control over our beliefs. While we cannot synchronically choose or decide to believe something, there are accounts of control that can be applied to beliefs and have beliefs come out as free. The debate over action control in the literature on moral responsibility demonstrates that voluntary control can be much more varied and nuanced than simply consciously deciding to act just prior to action. Oftentimes, we exercise control over our actions diachronically, over time. A sufficiently deep account of control allows for control over beliefs as well without synchronic choice or decision in a way that still grounds attributions of epistemic responsibility.

In the debate about moral responsibility, many philosophers have argued that a person can be responsible for an action even in circumstances where the person does not have access to alternative possibilities. John Martin Fischer and Mark Ravizza, for instance, distinguish between what they call regulative control from guidance control (Fischer and Ravizza, 1998). When a person exercises regulative control over an action, the person has the ability to act or not to act. Fischer and Ravizza are convinced that Harry Frankfurt's counterarguments show how such control is not a necessary condition for moral responsibility. Whether blame or praise is appropriate is not determined by whether the agent had options, but rather the actual sequence of events that led to the action. Against this, guidance control is a form of control that requires no access to alternative possibilities ("choice"). Instead, a person exercises guidance control over an action when the sequence that leads to the action is a result of a mechanism that is both reasons-responsive and owned by the agent.

One philosopher that has applied the account of Fischer and Ravizza on believing is McCormick (2011, 2015). McCormick argues that reasons-responsiveness applies very well to mechanisms that produce beliefs. A useful test for responsiveness is to imagine various counterfactual scenarios. Perception, for instance, is quite responsive to reasons. Let us say I believe there is a computer in front of me, because there is a computer in front of me. Let us also imagine what would happen, if that computer were taken away. Most likely I would cease to believe that there is a computer in front of me. If my belief that there is computer in front of me were a product of a drug-induced delusion, for instance, it would not be so responsive to perceptual evidence. So, an actual sequence of events that includes my normally functioning perceptual system is reasons-responsive to a much higher degree than, say, an actual sequence that includes drug-induced hallucinations.

For Fischer and Ravizza, reasons-responsiveness is not enough for responsibility. A person cannot be said to appropriately control her actions, if those actions are not issued by a mechanism that does not properly belong to the agent. The agent must take responsibility for the outputs of those mechanisms and they must be her own. The previous example of drugs causing a change in one's perception is an example where the mechanism is not the agent's own. So, the challenge is to demonstrate how an agent could own and take responsibility over her belief-producing mechanisms. McCormick thinks that this challenge can be met.

TAKING RESPONSIBILITY FOR ONE'S BELIEVING

Taking responsibility and owning one's belief-producing mechanisms are historical notions. First, I identify and recognize

[1]Peels (2013, 2017), for instance, rejects doxastic voluntarism but grounds epistemic responsibility practices on something close to believing. For (c), see Levy (2007). For important contributions to the debate, see Steup (2001). See also, Alston (2005), p. 58–81.

[2]For a version of Williams' argument, see Buckareff (2014).

the kinds of tendencies my epistemic faculties have had and also begin to understand their consequences. Second, this diachronic process also extends to the future: I begin to accept that I am being blamed and praised on the grounds of how my epistemic mechanisms meet given standards. Fischer and Ravizza take this as a process of building up one's identity over time (Fischer and Ravizza, 1998, p. 210–217). We train children to respond to blame and praise until they eventually internalize most of the instruction. They begin to feel appropriate emotional responses and accept that they are judged on the basis of their behavior.

According to McCormick, a similar process of ownership can take place with respect to our epistemic faculties (McCormick, 2011). She takes one reactive emotion, guilt, as an example. She argues that we sometimes feel guilty for having a belief. If this is indeed appropriate, it reveals that we implicitly take beliefs to be a result of our agency. She also examines various belief-producing mechanisms, like perception and memory. Perception is a standard example of a mechanism with respect to which our agency is completely passive. McCormick points out that this is not so. There are epistemic standards against which we measure our management of our perception. We must be mindful of the circumstances and whether we are under the influence of perception-impairing chemicals, like drugs. Again, we can distinguish between those cases where a person's belief is a product of a sequence gone haywire (drugs or psychotic hallucinations, bad environment, etc.) and between normally functioning sequences. A failure to do so is a failure of accepted epistemic standards. When a person becomes a member of an epistemic community and internalizes its norms, she accepts that she can be blamed and praised according to how she manages to meet these standards. While perceptual systems are not under direct voluntary control—a person cannot decide to believe—she, nevertheless, exercises some control over maintenance of her perceptual systems. For failures of this maintenance, she can be held accountable.

CONCLUSIONS

If the brief account I presented above is correct, it is indeed appropriate to describe and talk about believing as a dynamic process that involves our agency. On this account, believing does not simply happen to us but is a product of reason-responsive mechanisms that properly belong to us. Some of our beliefs are under our indirect control: we manage the cognitive mechanisms that issue them and control whether they are operate in the right environments. As members of an epistemic community, we have accepted that we are apt targets of epistemic blame and praise. If I manage my epistemic faculties poorly and adopt bad beliefs, I can be legitimately blamed for them.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

This paper was funded by Dr. Rüdiger Seitz, via the Volkswagen Foundation, Siemens Healthineers, and the Betz Foundation.

ACKNOWLEDGMENTS

The research for the article was graciously supported by the Finnish Cultural Foundation.

REFERENCES

Alston, W. (2005). *Beyond "Justification": Dimensions of Epistemic Evaluation*. Ithaca, NY: Cornell University Press. doi: 10.7591/9781501720574

Angel, H.-C., Oviedo, L., Paloutzian, R., Runehov, A., and Seitz, R. (eds.). (2017). *Processes of Believing: The Acquisition, Maintenance, and Change in Creditions*. Dordrecht: Springer. doi: 10.1007/978-3-319-50924-2

Audi, R. (2015). *Rational Belief*. Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780190221843.001.0001

Buckareff, A. (2014). "Deciding to believe redux" in *The Ethics of Belief*, eds J. Matheson and R. Vitz (Oxford: Oxford University Press) p. 33–50. doi: 10.1093/acprof:oso/9780199686520.003.0003

Fischer, J. M., and Ravizza, M. (1998). *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511814594

Levy, N. (2007). Doxastic responsibility. *Synthese* 155, 127–155. doi: 10.1007/s11229-005-3983-5

McCormick, M. (2011). Taking control of belief. *Philos. Explorat.* 14, 2. doi: 10.1080/13869795.2011.569745

McCormick, M. (2015). *Believing Against the Evidence.* New York, NY: Routledge. doi: 10.4324/9780203579146

Peels, R. (2013). Does doxastic responsibility entail the ability to believe otherwise? *Synthese* 190, 3651–3669. doi: 10.1007/s11229-012-0217-5

Peels, R. (2017). *Responsible Belief*. Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780190608118.001.0001

Steup, M. (ed.). (2001). *Knowledge, Truth, and Duty: Essays on Epistemic Justification, Responsibility, and Virtue*. Oxford: Oxford University Press.

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Visala. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Behavioral Neuroscience

OPINION
published: 19 April 2022
doi: 10.3389/fnbeh.2022.880504

Believing and Beliefs—Neurophysiological Underpinnings

*Rüdiger J. Seitz**

Department of Neurology, Centre of Neurology and Neuropsychiatry, LVR-Klinikum Düsseldorf, Medical Faculty, Heinrich-Heine-University Düsseldorf, Düsseldorf, Germany

Keywords: belief, fMRI, perception, valuation, language, representations, propositions

INTRODUCTION

The credition model posits that beliefs are the result of neural processes that involve the perception of external information and their valuation in terms of personal meaning determining a person's behavioral decisions (Seitz et al., 2018). These processes of believing typically evolve in a pre-linguistic fashion and include memory functions by which beliefs can be stored and recalled (Seitz et al., 2022). Thus, beliefs are fundamental representations of imaginative and emotional content that link an individual's prior experience with his/her future behavior. Importantly, people can become aware of what they believe and express it explicitly by "I believe …" (Oakley and Halligan, 2017; Seitz and Angel, 2020). Such propositions have a first-person perspective and can communicate the subject's certainty or trust into such a personally held belief to other people.

In this communication, the brain structures related to the processes of believing as identified by functional imaging are described. In the first part, imaging studies are presented in which healthy subjects processed statements of believing. The second part focuses on functional MRI studies addressing pre-linguistic processing involved in belief formation and updating.

VERBAL PROCESSING UNDERLYING BELIEVING

Secular political beliefs and religious beliefs are based on narratives that can be communicated by recital of stories or by written manifests. Ritual acts associated with these narratives lend emotional flavor to them by cognitive-emotional integration. Such beliefs correspond to so-called conceptual beliefs (**Figure 1A**). The first imaging study addressing the question which brain structures are involved in processing of a religious belief was by Azari et al. (2001). Christian Protestants were subjected to functional imaging while they recited Psalm 23. The strongest activation occurred in dorsal medial frontal cortex in comparison to reciting a nursery as well as to non-believing subjects (**Figures 1B,C**). Note, that in this study the neural representations of the Christian belief content was the research topic. This is different from the following three studies in which first-person assessments of believing were studied.

In one study, healthy subjects were required to indicate whether they agreed to statements about the involvement of God in the world such as "God protects one's life" (Kapogiannis et al., 2009). The pattern of activation involved also the dorsal medial frontal cortex besides a number of other cortical areas. It was suggested that the subjects engaged mentalizing processes to understand God's intent (Kapogiannis et al., 2009). A subsequent multivariate directional connectivity analysis showed that the religious subjects preferentially activated a pathway from inferolateral to dorsal medial frontal cortex. This pattern was interpreted as monitoring of the intent and involvement of supernatural agents. In contrast, perception of supranatural agents was found to engage pathways involved in fear regulation and affective mentalizing (Kapogiannis et al., 2014).

OPEN ACCESS

Edited by:
Nuno Sousa,
University of Minho, Portugal

Reviewed by:
Jesus Eduardo Pastor,
Princess University Hospital, Spain

Correspondence:
Rüdiger J. Seitz
seitz@neurologie.uni-duesseldorf.de

Specialty section:
This article was submitted to Emotion Regulation and Processing, a section of the journal Frontiers in Behavioral Neuroscience

Received: *21 February 2022*
Accepted: *23 March 2022*
Published: *19 April 2022*

Citation:
Seitz RJ (2022) Believing and Beliefs—Neurophysiological Underpinnings. Front. Behav. Neurosci. 16:880504. doi: 10.3389/fnbeh.2022.880504

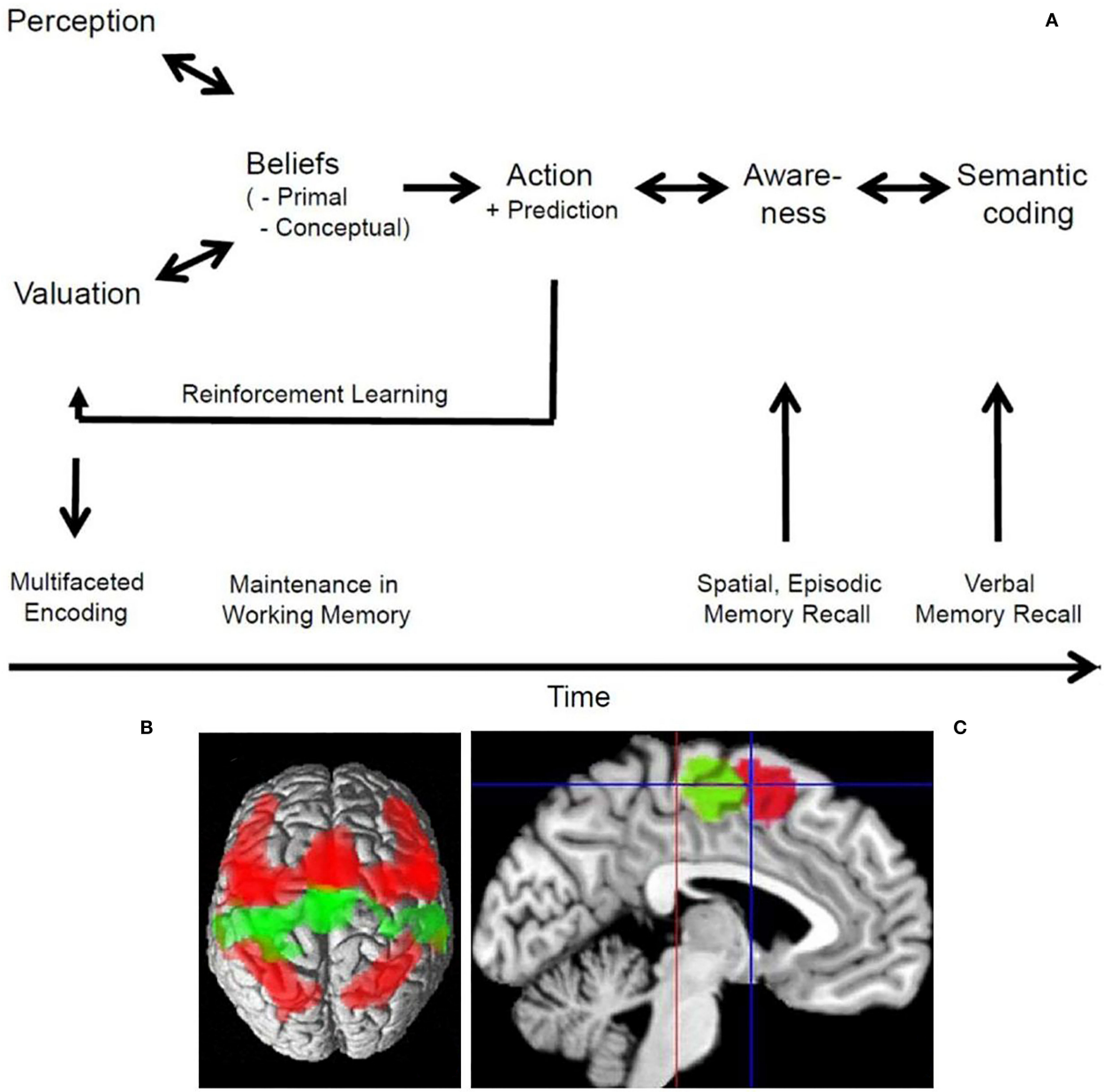


FIGURE 1 | The processes of believing and the dorsal medial frontal cortex. **(A)** Schematic drawing of the processes of believing and their close relation to memory functions. The bi-directional arrows account for putative bottom-up and top-down processing, while the uni-directional arrows indicate the major flow of information. Further details in Seitz et al. (2022). **(B)** The connectivity patterns of the SMA (green) and pre-SMA (red) as evident from a meta-analysis of functional imaging studies in a dorsal view on the brain. The SMA is predominantly connected bilaterally with the motor cortex, while the pre-SMA is predominantly connected bilaterally with the dorsolateral prefrontal as well as inferior and superior parietal cortex. Note, that the pre-SMA projects also to the SMA. **(C)** Medial aspect of the brain with cytoarchitectonic localization of the SMA (green) posterior and of the pre-SMA (red) anterior to the vertical plane through the anterior commissure (blue line); red line indicates the vertical plane through the posterior commissure. Further details in Ruan et al. (2018). To **(B,C)** the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/) applies.

In a more recent functional magnetic resonance imaging (fMRI) study healthy subjects were asked to decide whether propositions presented to them were true or false. These included statements that can be tested such as "I believe that hamsters are more common as pets than turtles". And there were statements that cannot be tested such as "I believe that giving love to others is the most important thing in my life". These assessments involved widespread, but non-overlapping cortical circuits (Howlett and Paulus, 2015). The dorsomedial prefrontal cortex, the precuneus and the cingulate gyrus were activated when the subjects were certain concerning their assessments of the testable statements, while the superior temporal gyrus was activated when the subjects

were certain concerning the non-testable statements. More recently, Chinese and Danish students were required to indicate in a yes-no response if they believed that adjectives presented to them described themselves, celebrities or had positive or negative valence. The behavioral data showed cultural group differences in self-construal, self-believing and celebrity-believing judgments. The fMRI data showed that there were common activations as well as significant differences across both groups of participants. Importantly, the dorsal medial frontal cortex was activated in the Chinese but not Danish students with regard to self-construal (Gao et al., 2022).

PRE-LINGUISTIC PROCESSES OF BELIEVING

The formation and updating of beliefs involve rapidly evolving neural processes such as perception, valuation, sensorimotor control, mentalizing, and perceptive-emotional integration. These are called primal beliefs or belief precursors and do not depend on language functions (Oakley and Halligan, 2017; Seitz and Angel, 2020). Conversely, people can state their primal or pre-linguistic beliefs verbally only after they have become aware of them. Inherent in these processes is the notion of the subjective first-person perspective of valuating of external information in terms of personal meaning and relevance. These representations have an imaginative character and are continuously updated by new information (**Figure 1A**). They build the basis on which subjects generate their spontaneous actions and make predictions of future events. These processes are maintained in putative parallel cortico-subcortical loops in the brain which was taken as basis for computational modeling of belief formation (Friston et al., 2017). From a methodological point of view the instructions to perform the tasks in functional imaging experiments were verbal statements. However, the neuropsychic processes initiated by them did not depend on language functions. Thus, the functional imaging studies addressed the question which structures of the human brain are engaged in relation to such pre-linguistic processes of believing. They are summarized here as follows.

The neural coding of emotional valence has been shown to involve widespread neural circuits distributed over different cortical and subcortical regions. The dorsolateral prefrontal cortex has been shown to be deeply interwoven with the integration of emotion and cognition (Gray et al., 2002; Okon-Singer et al., 2015). This also applies to affective and cognitive perspective taking (Healey and Grossman, 2018). Moreover, the prefrontal cortex and posterior cingulate direct attention to processes of unconscious threat (Etkin et al., 2009), while the right lateral prefrontal cortex was found to be involved in preference judgments (Elliott and Dolan, 1998). In addition, the basolateral amygdala and the nucleus accumbens are important brain structures related to the diversified aspects of valence encoding (Le Doux, 1996; Namburi et al., 2016; Vestergaard and Schultz, 2020). Likewise, it was found that a well-coordinated prefrontal-striatal network that is activated while a subject is experiencing a reward shapes preferences for future choices (Tanaka et al., 2020). Also, cognitive appraisal of emotions, belief updating, and self-perspective inhibition has been related to activity in a right fronto-parietal network (Miura et al., 2020). As a consequence, the lateral prefrontal cortex participates in the dynamic control of executive actions and in behavioral control (Mansuri et al., 2009). Importantly, positive and negative outcomes are encoded in the medial prefrontal cortex but with opposite signs in its ventral and dorsal subdivisions (Pischedda et al., 2020).

Besides its role in integrating cognitive and emotional information, the prefrontal cortex has been shown to be involved also in maintaining the concept of a personal self (Fossati et al., 2003). Specifically, activity in the dorsolateral prefrontal cortex was found in a phonological or semantic judgment task to be associated with priming effects (Lau and Passingham, 2007). Moreover, it was found that visually presented personally relevant words that signal important emotional clues engage a widely distributed set of brain regions including the dorsal medial and lateral prefrontal cortex (Huth et al., 2016). Further, emotion-denoting words were found to activate a large-scale neural network in the prefrontal cortex subserving the affective dimensions of valence and another network involving the left parahippocampus and dorsal anterior cingulate for affective arousal (Posner et al., 2009). Importantly, these processes did not activate language-related cortical areas.

Processing of events in the environment involves the dorsal cerebral midline structures including the supplementary motor area (SMA) and pre-SMA (**Figure 1C**). For example, when the cingulate is activated, it is likely that a negative event occurred (Jocham et al., 2009). This may be related to the time needed and effort invested to resolve a conflict (Kennerley et al., 2008; Mansuri et al., 2009). Also, during the generation and control of behavior, subliminal stimuli are thought to trigger inhibitory processes in extended prefrontal cortical areas that act on the pre-supplementary motor area (van Gaal et al., 2008). Notably, it has been found that anticipation of reward and punishment are mediated by opponent mechanisms but have some shared activations (Lake et al., 2019). Furthermore, activation of the orbitofrontal cortex reflects the subjective value of anticipated outcomes, whereas activation of the SMA reflects the probability of a persons' choice (FitzGerald et al., 2009). In contrast, activity in a cortico-subcortical network involving the striatum and the pre-SMA was found to be related to reward prediction (Hsu et al., 2009). Interestingly, involvement of the pre-SMA and bilaterally of the insula reflected subjective uncertainty (FitzGerald et al., 2009).

As humans develop subjective preferences and are able to make predictions about future events and other people's behavior, they need to decide what to do next, how to react to the actions of other people, and how to maximize the benefit between differential choices. Typically, these decisions can lead to either an immediate reward or to long-term satisfaction (Rolls, 2006). An interesting question is whether such choices require conscious awareness. Perceptual decisions have been found to be based on the matching of predicted and observed evidence in tests of perceptually ambiguous stimuli (Summerfield et al., 2006). Subjective preference judgments are mediated by the prefrontal cortex, medial orbitofrontal cortex, insula, and cingulate (Chaudhury et al., 2009). It was shown experimentally

that people make choices via the anterior prefrontal cortex using preferences of which they are not aware (Tusche et al., 2010). Similarly, day-to-day decisions were found to involve the ventromedial prefrontal related to valuation and choice (Levy and Glimcher, 2012; Kumar et al., 2019; Koscik et al., 2020). Decisions concerning reasoning about other peoples face expressions were shown to be made with high accuracy in a time window too little to account for conscious awareness (Prochnow et al., 2013). Nevertheless, the entire cortical processing network related to emotional face perception was involved. In contrast, the fusiform face area was more active during supraliminal face presentation. This corresponded to the observations that brain regions, including the amygdala, become activated by emotional faces only when sufficient attentional resources concerning the effects of valence are available (Pessoa et al., 2002). Interestingly, observing people interacting with each other activated the posterior superior temporal cortex related to meta-theoretical inference about what is being observed (Isik et al., 2017). It is of note that the pre-SMA was involved in such decisions (Prochnow et al., 2014). Therefore, it can be argued that the pre-SMA integrates online information processing in the dorsolateral prefrontal cortex with motor command processing (**Figure 1B**). This is consistent with the observation that preference adjustments in difficult decisions are related to activity in a widespread left dorsolateral prefrontal-midparietal network (Voigt et al., 2019). Such, findings support the view that computation of the expected value in mesolimbic structures represents an affective component, whereas cortical regions represent a probabilistic component, and may integrate the two (Knutson et al., 2005).

DISCUSSION

Beliefs are pre-linguistic representations of imaginative and emotional content that link an individual's prior experience with his/her future behavior. These functions enable humans to infer social meaning from other people's behavior and to make corresponding attributions (Malle and Korman, 2022). Furthermore, humans can become aware of their beliefs and express their content in the form of semantic expressions. It was shown here that processing of beliefs engages widespread cortical circuits related to inferential attribution, cognitive-emotional integration, and language functions. The dorsal medial frontal cortex comprising the so-called pre-SMA was shown to be a critical hub with a large-scale cortico-subcortical loop involving the thalamus and reciprocal connectivity to prefrontal and parietal cortical areas (Reid et al., 2015; Ruan et al., 2018). The overlap of this connectivity pattern with the cortical circuitry related to working memory and the so-called default network (Reid et al., 2015) accords with a prominent role also in belief evaluation (Sugiura et al., 2015). Belief evaluation is a language-based function by which humans can consider critically what they believe and how this corresponds to their predictions (Coltheart et al., 2011). Conversely, patients with neurological and psychiatric diseases provided evidence that focal brain lesions can interfere with the formation, updating and evaluation of beliefs (Coltheart et al., 2011; Seitz, 2022). Thus, brain diseases interfering with the processes of believing can induce abnormal beliefs that can cause deviant behavior.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

This study was funded by the Volkswagen Foundation: #99835, Siemens Healthineers, and the Betz Foundation. Siemens Healthineers were not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

REFERENCES

Azari, N. P., Nickel, J., Wunderlich, G., Niedeggen, M., Hefter, H., Tellmann, L., et al. (2001). Neural correlates of religious experience. *Eur. J. Neurosci.* 13, 1649–1652. doi: 10.1046/j.0953-816x.2001.01527.x

Chaudhury, A. M., Parkinson, J. A., Hinton, E. C., Owen, A. M., and Roberts, A. C. (2009). Preference judgements involve a network of structures within frontal, cingulate and insula cortices. *Eur. J. Neurosci.* 29, 1047–1055. doi: 10.1111/j.1460-9568.2009.06646.x

Coltheart, M., Langdon, R., and McKay, R. (2011). Delusional belief. *Ann. Rev. Psychol.* 62, 271–298. doi: 10.1146/annurev.psych.121208.131622

Elliott, R., and Dolan, R. J. (1998). Neural response during preference and memory judgments for subliminally presented stimuli: a functional neuroimaging study. *J. Neurosci.* 18, 4697–4704. doi: 10.1523/JNEUROSCI.18-12-04697.1998

Etkin, A., Pittenger, C., Polan, H. J., and Kandel, E. R. (2009). Toward a neurobiology of psychotherapy: basic science and clinical applications. *J. Neuropsychol.* 17, 145–158. doi: 10.1176/jnp.17.2.145

FitzGerald, T. H. B., Seymour, B., and Dolan, R. J. (2009). The role of human orbitofrontal cortex in value comparions for incommensurable objects. *J. Neurosci.* 29, 8388–8395 doi: 10.1523/JNEUROSCI.0717-09.2009

Fossati, P., Hevenor, S. J., Graham, S. J., Grady, C., Keightley, M. L., Craik, F., et al. (2003). In search of the emotional self: an fMRI study using positive and negative emotional words. *Am. J. Psychiatry* 160, 1938–1945. doi: 10.1176/appi.ajp.160.11.1938

Friston, K. J., Parr, T., and de Vries, B. (2017). The graphical brain: belief propagation and active inference. *Network Neurosci.* 1, 381–414. doi: 10.1162/NETN_a_00018

Gao, T., Han, X., Bang, D., and Han, S. (2022). Cultural differences in neurocognitive mechanisms of believing. *Neuroimage* 250, 118954. doi.org/10.1016/j.neuroimage.22.118954

Gray, J. R., Braver, T. S., and Raichle, M. E. (2002). Integration of emotion and cognition in the lateral prefrontal cortex. *Proc. Natl. Acad. Sci. USA* 99, 4115–4120. doi: 10.1073/pnas.062381899

Healey, M. L., and Grossman, M. (2018). Cognitive and affective perspective-taking: evidence for shared and dissociable anatomical substrates. *Front. Neurol.* 9, 491. doi: 10.3389/fneur.2018.00491

Howlett, J. R., and Paulus, M. P. (2015). The neural basis of testable and non-testable beliefs. *PLoS One* 10, e0124596; doi: 10.1371/journal.pone.0124596

Hsu, M., Krajbich, I., Zhao, C., and Camerer, C. F. (2009). Neural response to reward anticipation under risk is nonlinearin probabilities. *J. Neurosci.* 29, 2231–2237. doi: 10.1523/JNEUROSCI.5296-08.2009

Huth, A. G., de Heer, W. A., Griffiths, T. L., and Theunissen, F. E. (2016). Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature* 532, 453–458. doi: 10.1038/nature17637

Isik, L., Koldewyn, K., Beeler, D., and Kanwisher, N. (2017). Perceiving social interactions in the posterior superior temporal sulcus. *Proc. Natl. Acad. Sci. USA* 114, E9145–E9152. doi: 10.1073/pnas.1714471114

Jocham, G., Neuman, J., Klein, T. A., Danielmeier, C., and Ullsperger, M. (2009). Adaptive coding of action values in human rostral cingulate zone. *J. Neurosci.* 29, 7489–7496. doi: 10.1523/JNEUROSCI.0349-09.2009

Kapogiannis, D., Barbey, A. K., Su, M., Zamboni, G., Krueger, F., and Grafman, J. (2009). Cognitive and neural foundations of religious belief. *Proc. Natl. Acad. Sci. USA*. 106, 4876–4881 doi: 10.1073/pnas.0811717106

Kapogiannis, D., Deshpande, G., Krueger, F., Thornburg, M. P., and Grafman, J. H. (2014). Brain networks shaping religious belief. *Brain Connect.* 4, 70–79. doi: 10.1089/brain.2013.0172

Kennerley, S. W., Dahmubed, A. F., Lara, A. H., and Wallis, J. D. (2008). Neurons in the frontal lobe encode value of multiple decision variables. *J. Cogn. Neurosci.* 21, 1162–1178. doi: 10.1162/jocn.2009.21100

Knutson, B., Taylor, J., Kaufman, M., Peterson, R., and Glover, G. (2005). Distributed neural representation of expected value. *J. Neurosci.* 25, 4806–4812. doi: 10.1523/JNEUROSCI.0642-05.2005

Koscik, T. R., Man, V., Jahn, A., Lee, C. H., and Cunningham, W. A. (2020). Decomposing the neural pathways in a simple, value-based choice. *Neuroimage* 214, 116764. doi: 10.1016/j.neuroimage.2020.116764

Kumar, P., Pisconi, A., Bondy, E., Kremens, R., Singleton, P., Pizzagalli, D. A., et al. (2019). Delineating the social valuation network in adolescents. *Soc. Cogn. Affect. Neurosci*. 14, 1159–1166. doi 10.1093/scan/nsz086

Lake, J. I., Spielberg, J. M., Infantolino, Z. P., Crocker, L. D., Yee, C. M., Holler, W., et al. (2019). Reward anticipation and punishment anticipation are instantiated in the brain via opponent mechanisms. *Psychopharmacology* 56, e13381. doi: 10.1111/psyp.13381

Lau, H. C., and Passingham, R. E. (2007). Unconcious activation of the cognitive control system in the human prefrontal cortex. *J. Neurosci.* 27, 5805–5811. doi: 10.1523/JNEUROSCI.4335-06.2007

Le Doux, J. E. (1996). The Emotional Brain. The Mysterious Underpinnings of Emotional Live. New York, NY: Simon and Schuster.

Levy, D. J., and Glimcher, P. W. (2012). The root of all value: a neural common currency for choice. *Curr. Opin. Neurobiol.* 22, 1027–1038. doi: 10.1016/j.conb.2012.06.001

Malle, B. F., and Korman, J. (2022). "Attribution theory," in *Wiley-Blackwell Encyclopedia of Social Theory*, eds B. S. Turner. Malden, MA: Wiley Blackwell (in press).

Mansuri, F. A., Tanak, K., and Buckley, M. J. (2009). Conflict-induced behavioural adjustment: a clue to the executive functions of the prefrontal cortex. *Nat. Rev.* 10, 141–152. doi: 10.1038/nrn2538

Miura, N., Suguira, M., Nozawa, T., Yamamoto, Y., Sasaki, Y., Hamamoto, Y., et al. (2020). Taking another's perspective promotes right parieto-frontal activity that reflects open-minded thought. *Soc. Neurosci*. 15, 282–295. doi: 10.1080/17470919.2019.1710249

Namburi, P., Al-Hasani, R., Calhoon, G. G., Bruchas, M. R., and Tye, K. M. (2016). Architecural representation of valuence in the limbic system. *Neuropsychopharmacology* 41, 1697–1715. doi: 10.1038/npp.2015.358

Oakley, D. A., and Halligan, P. W. (2017). Chasing the rainbow: the non-conscious nature of being. *Front. Psychol.* 8, 1924. doi: 10.3389/fpsyg.2017.01924

Okon-Singer, H., Hendler, T., Pessoa, L., and Shackman, A. J. (2015). The neurobiology of emotion-cognition interactions: fundamental questions and strategies for future research. *Front. Hum. Neurosci.* 9, 58. doi: 10.3389/fnhum.2015.00058

Pessoa, L., McKenna, M., Gutierrez, E., and Ungerleider, L. G. (2002). Neural processing of emotional faces requires attention. *Proc. Natl. Acad. Sci. USA* 99, 11458–11463. doi: 10.1073/pnas.172403899

Pischedda, D., Palminteri, S., and Coricelli, G. (2020). The effect of counterfactual information on outcome value coding in medial prefrontal and cingulate cortex: from an absolute to a relative neural code. *J. Neurosci.* 40, 3268–3277. doi: 10.1523/JNEUROSCI.1712-19.2020

Posner, J., Russell, J. A., Gerber, A., Gorman, D., Colibazzi, T., Yu, S., et al. (2009). The neurophysiological bases of emotion: an fMRI study of the affective circumplex using emotion-denoting words. *Hum. Brain Mapp.* 30, 883–895. doi: 10.1002/hbm.20553

Prochnow, D., Brunheim, S., Kossack, H., Eickhoff, S., Markowitsch, H. J., and Seitz, R. J. (2014). Anterior and posterior subareas of the dorsal lateral prefrontal cortex in socially relevant decisions based on masked affect expressions. *F1000research*. 3, 212. doi 10.12688/f1000research.4734.1

Prochnow, D., Kossack, H., Brunheim, S., Müller, K., Wittsack, H.-J., Markowitsch, H.-J., et al. (2013). Processing of subliminal facial expressions of emotion: a behavioral and fMRI study. *Soc. Neurosci.* 8, 448–61. doi: 10.1080/17470919.2013.812536

Reid, A. T., Bzdok, D., Langner, R., Fox, P. T., Laird, A. R., Amunts, K., et al. (2015). Multimodal connectivity mapping of the human left anterior and posterior lateral prefrontal cortex. *Brain Struct. Funct*. 221, 2589–2605. doi: 10.1007/s00429-015-1060-5

Rolls, E. T. (2006). Brain mechanisms underlying flavour and appetite. *Philos. Trans. R. Soc. London B Biol. Sci.* 361, 1123–1136. doi: 10.1098/rstb.2006.1852

Ruan, J., Bludau, S., Palomero-Gallagher, N., Caspers, S., Mohlberg, H., Eickhoff, S. B., et al. (2018). Cytoarchitecture, probability maps, and functions of the human supplementary and pre-supplementary motor areas. *Brain Struct. Funct*. 223, 4169–4186. doi: 10.1007/s00429-018-1738-6

Seitz, R. J., Angel, H., and Paloutzian, R. F. (2022). Bridging the gap between believing and memory functions. *Europe's J. Psychol. Psych. Arch.* doi: 10.23668/psycharchives.5421. [Epub ahead of print].

Seitz, R. J. (2022). Beliefs—a challenge in neuropsychological disorders. *J. Neuropsychol.* 16: 21–37 doi: 10.1111/jnp.12249

Seitz, R. J., Angel (2020). Belief formation – a driving force for brain evolution. *Brain Cogn* 140, 105548. doi: 10.1016/j.bandc.2020.105548

Seitz, R. J., Angel, H.-F., and Paloutzian, R. F. (2018). From believing to belief: a general theoretical model. *J. Cogn. Neurosci.* 30, 1254–1264. doi: 10.1162/jocn_a_01292

Sugiura, M., Seitz, R. J., and Angel, H.-F. (2015). Models and neural bases of the believing process. *J. Behav. Brain Sci.* 5, 12–23. doi: 10.4236/jbbs.2015.51002

Summerfield, C., Egner, T., Greene, M., Koechlin, E., Mangels, J., and Hirsch, J. (2006). Predictive codes for forthcoming perception in the frontal cortex. *Science* 314, 1311–1314. doi: 10.1126/science.1132028

Tanaka, D., Aoki, R., Suzuki, S., Takeda, M., Nakahara, K., and Jimura, K. (2020). Self-controlled choice arises from dynamic prefrontal signals that enable future anticipation. *J. Neurosci.* 40, 9736–9750. doi: 10.1523/JNEUROSCI.1702-20.2020

Tusche, A., Bode, S., and Haynes, J. D. (2010). Neural responses to unattended products predict later consumer choices. *J. Neurosci.* 30, 8024–8031. doi: 10.1523/JNEUROSCI.0064-10.2010

van Gaal, S., Ridderinkhof, K. R., Fahrenfort, J. J., Schulte, H. S., and Lamme, V. A. F. (2008). Frontal cortex mediates unconsciously triggered inhibitory control. *J. Neurosci.* 28, 8053–8062. doi: 10.1523/JNEUROSCI.1278-08.2008

Vestergaard, M. D., and Schultz, W. (2020). Retrospective valuation of experienced outcome encoded in distinct reward representations in the anterior insula and amygdala. *J. Neurosci.* 40, 8938–8950. doi: 10.1523/JNEUROSCI.2130-19.2020

Voigt, K., Murawski, C., Speer, S., and Bode, S. (2019). Hard decisions shape the neural coding of preferences. *J. Neurosci.* 39, 718–726. doi: 10.1523/JNEUROSCI.1681-18.2018

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Seitz. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Behavioral Neuroscience

TYPE Opinion
PUBLISHED 11 January 2023
DOI 10.3389/fnbeh.2022.951974

Check for updates

OPEN ACCESS

EDITED BY
Sushil K. Jha,
Jawaharlal Nehru University, India

REVIEWED BY
Shakeel Mohamed,
Technical University of Malaysia Malacca, Malaysia

*CORRESPONDENCE
Raimund Kleiser
raimund.kleiser@kepleruniklinikum.at

SPECIALTY SECTION
This article was submitted to Emotion Regulation and Processing, a section of the journal Frontiers in Behavioral Neuroscience

RECEIVED 24 May 2022
ACCEPTED 21 December 2022
PUBLISHED 11 January 2023

CITATION
Sonnberger M, Widmann C, Potthoff D, Seitz RJ and Kleiser R (2023) Emotion recognition in evolving facial expressions: A matter of believing. *Front. Behav. Neurosci.* 16:951974. doi: 10.3389/fnbeh.2022.951974

COPYRIGHT
© 2023 Sonnberger, Widmann, Potthoff, Seitz and Kleiser. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Emotion recognition in evolving facial expressions: A matter of believing

Michael Sonnberger[1], Carola Widmann[1], Denise Potthoff[2], Rüdiger J. Seitz[2] and Raimund Kleiser[1]*

[1]Institute of Neuroradiology, Kepler-University Linz, Linz, Austria, [2]Department of Neurology, Center of Neurology and Neuropsychiatry, LVR-Klinikum Düsseldorf, Medical Faculty, Heinrich-Heine University Düsseldorf, Düsseldorf, Germany

KEYWORDS

emotion, empathy, beliefs, face expression, fMRI

Introduction

Since the enlightenment period, beliefs have been considered widely as incompatible with science. This has promoted a reservation toward the notion of beliefs in the contemporary Western societies and the natural sciences, in particular. More recently, however, people have become aware that belief formation and believing can be a topic of increasing interest for a scientific discourse. Probably, this may have resulted from the observation that religious beliefs appeared as motivation for initiating outbursts of violence. It is important to realize, however, that beliefs are not limited to religious and political beliefs that are based on the narratives, but also comprise the so-called primal beliefs that do not depend on language functions as they concern objects and events in the environment (Seitz, 2022).

Believing is composed of cerebral processes involving the perception of external information and spontaneous appraisal of that information in terms of subjective value or meaning (Seitz et al., 2018). An important type of external information is human face, because facial expressions are considered as a human capacity to convey the emotional state of the given person (Russell, 1994). While the previous research on facial expressions of emotion has focused on the study of six basic categories, e.g., happiness, surprise, anger, sadness, fear, and disgust, recently, more than 20 compound facial expressions of emotions were identified that both can be produced and recognized as well (Du et al., 2014). In fact, humans are highly skilled to recognize the emotions in the rapidly changing facial expressions of other people (Fiske et al., 2007). Recognition of faces and facial expressions can be impaired in psychopathic disorders and alexithymia (Kyranides et al., 2022) as well as by the face masks that cover the nose and mouth (Kleiser et al., 2022). Importantly, however, the observing subject believes that she/he has recognized the facial expression of the other person and trust this belief (Brashier and Marsh, 2020). Moreover, upon recognition of the emotion in the facial expression of the other person, the facial muscles of the observing subject change in a corresponding fashion. This phenomenon demonstrated by electromyographic recordings was called facial mimicry (Franz et al., 2021). Accordingly, believing has an immediate impact on the expressive behavior of the believing subject (Seitz et al., 2022). The more pronounced the facial expressions are, the more likely the observing subject has recognized the observed emotion correctly and the more certain can she/he be in that belief.

Here, we requested healthy subjects to recognize the emotions in facial expressions that were displayed to them in video clips. Because we were interested in determining when the subjects believed to have recognized the emotions, we used video clips in which the emotional face expressions evolved within 20 s out of a neutral face. This allowed us to analyze the process of emotion recognition as compared to viewing the neutral face expressions and empathizing with the emotion seen in the face in the video clip. Empathy is the ability to take the other person's perspective that is considered a key element in entertaining interpersonal relationships (Bird and Viding, 2014). In a functional magnetic resonance imaging study, we found that recognition of an emotion in another person's facial expression results in the activation of large-scale cortico-subcortical circuits related to visual perception, emotion regulation, and action generation.

Emotion recognition

Overall, 16 healthy subjects (8 females, 8 males, 25 ± 6 years, normal or corrected-to-normal vision) passed a screening for alexithymia (TAS-20, Bagby et al., 1994) and capability of empathy (SPF, http://psydok.sulb.unisaarland.de/volltexte/2009/2363/pdf/SPF_Artikel.pdf). They gave informed written consent to participate in the study that was approved by the local ethics committee and conducted according to the Declaration of Helsinki. Male and female facial expressions of happiness, sadness, fear, and anger consisted of depersonalized frontal black and white images (Averaged Karolinska Institute AKDEF). Each emotion starting from a neutral facial expression evolving over time up to the strongest expression of the emotion (30 images of 750 ms each) was presented. The subjects were instructed to press a button as soon as they recognized the emotion or felt that they empathized with the emotional expressions.

On average, the subjects recognized the emotion anger, fear, and sadness when each emotion had evolved in the video clips to ∼80% (Figure 1). For comparison, happiness was recognized already when the emotion had evolved to some 40%, which is in accordance with the other studies (Adolphs, 2002). Interestingly, empathizing occurred while the facial expressions were still evolving with a similar delay across the four basic emotions (Figure 1).

Brain activation

The focus of this functional magnetic resonance imaging (fMRI) study was to map the brain regions related to the processes of emotion recognition in stereotactic space (Talairach and Tournoux, 1988). As validated in the brain of human primates, the changes in oxygenated blood as measured with fMRI are temporally and spatially related to the electrical field potential changes in neuronal assemblies following a definite sensory stimulus (Logothetis et al., 2001). fMRI was performed with a 3T MRI scanner (Siemens Magnetom Skyra) while the subjects were lying comfortably and viewed a mirror above them. Through this mirror, they observed the video clips. With a stimulation time of 22,500 ms and a control condition of 10,500 ms, this resulted in a total of 33,000 ms per block. A fixation cross was used as control stimulus shown between each trial to reset the BOLD signal (control condition). Each emotion was repeated six times, multiplied by four emotions in the two sexes resulting in a total of 48 repetitions and thus a total measurement time of 26.4 min. Whole-brain image analysis was done using the Brainvoyager QX software package version 21.4 (Brain Innovation, Maastricht, the Netherlands) as detailed elsewhere (Kleiser et al., 2017, 2022).

At that point when the subjects indicated by pressing a button to have recognized the emotions, there was a strong and widespread activation pattern involving cortical and subcortical brain structures (Figure 1). For comparison, before this time point, the activations were of a far weaker intensity ($p<0.05$; FDR-corrected) occurring in brain areas related to the ventral pathway for the processing of shape, color, and faces, extending to V4, and to the fusiform gyrus (Courtney and Ungerleider, 1997). In addition, the dorsal pathway, related to the processing of motion such as visual area V5, the lateral parietal cortex, and the frontal eye fields were involved. In the phase, when the subjects indicated empathy with the emotions, there were activations ($p < 0.05$; FDR-corrected) as known to be activated in empathic processing such as the inferior frontal gyrus and the superior temporal gyrus (Shamay-Tsoory et al., 2009; Schurz et al., 2017). It is well possible that the subjects differed in how intensively and how long they were able to maintain empathy.

Importantly, however, the activations at the time point of recognition of the emotions exceeded an even higher level of significance ($p<0.01$; FDR-corrected). In the cerebral cortex, they involved the dorsomedial frontal cortex including the supplementary motor area (SMA) and pre-SMA, the dorsolateral frontal cortex, the inferior frontal cortex, and the inferior temporal cortex in a virtually symmetric pattern in both cerebral hemispheres (Figure 1). In addition, there were strong activations in the basal ganglia, the entire thalamus, and amygdala and midbrain nuclei including the red nucleus. Finally, there was an involvement in the cerebellar vermis. These activations are compatible with the notion of an engagement of parallel cortico-subcortical cerebellar circuits (DeLong et al., 1984).

Discussion

We have shown in the healthy subjects that believing to recognize the emotions in facial expressions engaged a widespread and distinctive network involving occipital, parietal, and frontal cortical areas. These areas are known to participate

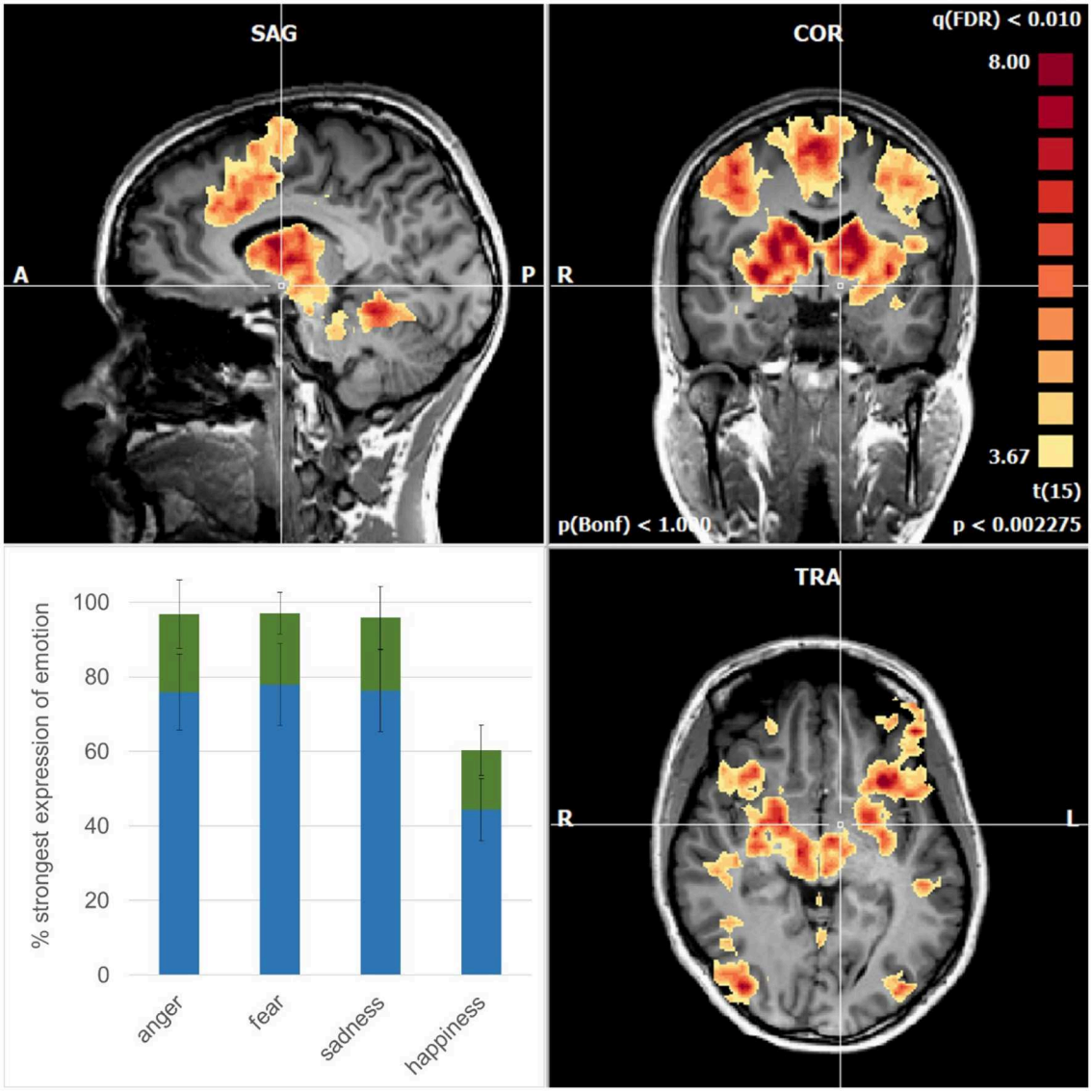


FIGURE 1
Mean activations are related to the recognition of the emotions in sagittal **(upper left)**, coronal **(upper right)**, and axial planes **(lower right)**. Note the symmetric pattern involves cortical areas and subcortical structures such as the basal ganglia, thalamus, amygdala, and brainstem nuclei. The cross of the stereotactic coordinates (x −4, y −4, z −6) signifies the left hypothalamus that is spared. **(lower left)** Degree of evolution of the emotional face expressions when the subjects recognized (blue) and empathized (green) with the emotions; error bars: standard deviations.

in the face recognition (Xu et al., 2021), oculomotor control (Pierrot-Deseilligny et al., 2004), and imitation of movement (Heiser et al., 2003). Moreover, strong activity was found also in subcortical structures such as the basal ganglia and the thalamus as a part of a highly developed cortico-subcortical relay circuitry, supporting these functions (DeLong et al., 1984). Furthermore, activity was found in the amygdala—a central neural component in emotion processing (Packard et al., 2021). Notably, this widespread pattern of cortical, subcortical, and cerebellar activations was similar to that recently observed in 944 participants during memory encoding of emotional pictures (Fastenrath et al., 2022). It is tempting to speculate that such an extended pattern of enhanced brain activity reflects the complexity of cerebral processing that may be suited to afford human conscious awareness (Greenfield and Collins, 2005).

With our experimental design, we were able to expand the duration of face presentation before the subjects recognized the emotion in the video clips. This allowed us to analyze the process of believing and to determine when the emotions were recognized correctly. It was amazing that this point occurred after the emotions were expressed to some 80%, and only happiness was recognized far earlier, corresponding to similar

findings by others (Adolphs, 2002). The belief that the emotions were recognized correctly was substantiated subsequently when emotion was more pronounced. Thus, the subjects' trust in their correct recognition of the emotion was confirmed instantaneously in a rebound manner. This aspect probably also contributed to the strength of the activation pattern observed. In fact, these activations were far stronger than those related to viewing the faces that appeared neutral while empathizing with the emotional face expressions. Nevertheless, it was possibly the imaging that correlates with oscillatory binding of brain activity, when subjects become aware of information processing (Engel and Singer, 2001). That empathizing occurred only shortly later, suggesting that the awareness of the emotion was the bottle-neck process preceding empathizing. Moreover, subjective reports of the subjects support that the assumption empathizing with the emotions was strengthened by the slow progression into the emotion, as compared to an immediate exposure to an outspoken emotion upon viewing the static images of facial expressions.

People process sensory information with ease, which makes them susceptible to trusting these perceptions (Brashier and Marsh, 2020). Concerning affect recognition children, in contrast to adults, have been reported to observe both the eyes and the mouth (Guarnera et al., 2018). As happy faces typically have an open mouth that uncovers the teeth, this may serve as a clue for the observer to identify a happy emotion faster as compared to the other emotional states. The other basic emotions, such as sadness, anger, and fear, were recognized with similar ease (Kleiser et al., 2022). Importantly, however, humans believe that their perceptions are true reflections of the emotional states of the persons in their environment. This enables them to streamline the multitude of their sensory sensations according to those that are subjectively relevant for them and to select their subsequent behavioral actions accordingly (Seitz et al., 2022). Consequently, believing has not only a perceptive aspect about the subject's past experience, but also a prospective aspect concerning decision-making regarding the alternative actions with associated the predictions of what these actions will lead to and how the environment may react to these actions. The findings presented here provide empirical evidence for a putative neural basis for such processes of believing that afford intuitive, prelinguistic action generation (Seitz, 2022). Ultimately, they are apparently suited to support the concept that believing is a fundamental brain function (Angel and Seitz, 2016).

Author contributions

RK and DP contributed to conception and design of the study. RK and CW carried out the measurements. RK, CW, and MS performed data post-processing and statistical analysis. RS, RK, and MS wrote sections of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

Funding

This paper was funded by RS, *via* the Volkswagen Foundation, Siemens Healthineers, and the Betz Foundation. The funders were not involved in the study design, collection, analysis, interpretation of data, the writing of this article, or the decision to submit it for publication.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Adolphs, R. (2002). Neural systems for recognizing emotion. *Curr. Opin. Neurobiol.* 12, 169–177. doi: 10.1016/S0959-4388(02)00301-X

Angel, H. F., and Seitz, R. J. (2016). Process of believing as fundamental brain function: the concept of credition. *SFU Res. Bull.* 3, 1–20. doi: 10.15135/2016.4.1.1-20

Bagby, R. M., Parker, J. D., and Taylor, G. J. (1994). The twenty-item Toronto Alexithymia Scale-I. Item selection and cross-validation of the factor structure. *J. Psychosom. Res.* 38, 23–32. doi: 10.1016/0022-3999(94)90005-1

Bird, G., and Viding, E. (2014). The self to other model of empathy: providing a new framework for understanding empathy impairments in psychopathy, autism, and alexithymia. *Neurosci. Biobehav. Rev.* 47, 520–532. doi: 10.1016/j.neubiorev.2014.09.021

Brashier, N. M., and Marsh, E. J. (2020). Judging truth. *Ann. Rev. Psychol.* 71, 499–515. doi: 10.1146/annurev-psych-010419-050807

Courtney, S. M., and Ungerleider, L. G. (1997). What fMRI has taught us about human vision. *Curr. Opin. Neurobiol.* 7, 554–561. doi: 10.1016/S0959-4388(97)80036-0

DeLong, M. R., Alexander, G. E., Georgopoulos, A. P., Crutcher, M. D., Mitchell, S. J., Richardson, R. T., et al. (1984). Role of basal ganglia in limb movements. *Hum. Neurobiol.* 2, 235–244.

Du, S., Tao, Y., and Martinez, A. M. (2014). Compound facial expressions of emotion. *Proc. Natl. Acad. Sci. USA* 111, E1454–1462. doi: 10.1073/pnas.1322355111

Engel, A. K., and Singer, W. (2001). Temporal binding and the neural correlates of sensory awareness. *Trends Cogn. Sci.* 5, 16–25. doi: 10.1016/S1364-6613(00)01568-0

Fastenrath, M., Spalek, K., Coynel, D., Loos, E., Milnik, A., Egli, T., et al. (2022). Human cerebellum and corticocerebellar connections involved in emotional memory enhancement. *Proc. Natl. Acad. Sci. USA* 119, e2204900119. doi: 10.1073/pnas.2204900119

Fiske, S. T., Cuddy, A. J., and Glick, P. (2007). Universal dimensions of social cognition: warmth and competence. *Trends Cogn. Sci.* 11, 77–83. doi: 10.1016/j.tics.2006.11.005

Franz, M., Nordmann, M. A., Rehagel, C., Schäfer, R., Müller, T., Lundqvist, D., et al. (2021). It is in your face—alexithymia impairs facial mimicry. *Emotion*. 21, 1537–1549. doi: 10.1037/emo0001002

Greenfield, S. A., and Collins, T. F. T. (2005). A neuroscientific approach to consciousness. *Prog. Brain Res.* 150, 11–23. doi: 10.1016/S0079-6123(05)50002-5

Guarnera, M., Magnano, P., Pellerone, M., Cascio, M. I., Squatrito, V., Buccheri, S. L., et al. (2018). Facial expressions and the ability to recognize emotions from the eyes or mouth: a comparison among old adults, young adults, and children. *J. Genetic Psychol.* 179, 297–310. doi: 10.1080/00221325.2018.1509200

Heiser, M., Iacoboni, M., Maeda, F., Marcus, J., and Mazziotta, J. C. (2003). The essential role of Broca's area in imitation. *Eur. J. Neurosci.* 17, 1123–1128. doi: 10.1046/j.1460-9568.2003.02530.x

Kleiser, R., Raffelsberger, T., Trenkler, J., Meckel, S., and Seitz, R. J. (2022). What influence do face masks have on reading emotions in faces? *NeuroImage Rep.* 2, 100141. doi: 10.1016/j.ynirp.2022.100141

Kleiser, R., Stadler, C., Wimmer, S., Matyas, T., and Seitz, R. J. (2017). An fMRI study of training voluntary smooth circular eye movements. *Exp. Brain Res.* 235, 819–831. doi: 10.1007/s00221-016-4843-x

Kyranides, M. N., Christofides, D., and Çetin, M. (2022). Difficulties in facial emotion recognition: taking psychopathic and alexithymic traits into account. *BMC Psychol.* 10, 239. doi: 10.1186/s40359-022-00946-x

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., and Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature* 412, 150–157. doi: 10.1038/35084005

Packard, M. G., Gadberry, T., and Goodman, J. (2021). Neural systems and the emotion-memory link. *Neurobiol. Learn Mem.* 185, 107503. doi: 10.1016/j.nlm.2021.107503

Pierrot-Deseilligny, C., Milea, D., and Müri, R. M. (2004). Eye movement control by the cerebral cortex. *Curr. Opin. Neurol.* 17, 17–25. doi: 10.1097/00019052-200402000-00005

Russell, J. A. (1994). Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychol. Bull.* 115, 102–141. doi: 10.1037/0033-2909.115.1.102

Schurz, M., Tholen, M. G., Perner, J., Mars, R. B., and Sallet, J. (2017). Specifying the brain anatomy underlying temporo-parietal junction activations for theory of mind: a review using probabilistic atlases from different imaging modalities. *Hum. Brain Mapp.* 38, 4788–4805. doi: 10.1002/hbm.23675

Seitz, R. J. (2022). Believing and beliefs—neurophysiological underpinnings. *Front. Behav. Neurosci.* 16, 880504. doi: 10.3389/fnbeh.2022.880504

Seitz, R. J., Angel, H. F., Paloutzian, R. F., and Taves, A. (2022). Believing and social interactions: effects on bodily expressions and personal narratives. *Front. Behav. Neurosci.* 16, 894219. doi: 10.3389/fnbeh.2022.894219

Seitz, R. J., Paloutzian, R. F., and Angel, H. F. (2018). From believing to belief: a general theoretical model. *J. Cogn. Neurosci.* 30, 1254–1264. doi: 10.1162/jocn_a_01292

Shamay-Tsoory, S., Aharon-Peretz, J., and Perry, D. (2009). Two systems for empathy: a double dissociation between emotional and cognitive empathy in inferior frontal gyrus versus ventromedial prefrontal lesions. *Brain* 132, 617–627. doi: 10.1093/brain/awn279

Talarairach, J., and Tournoux, P. (1988). *Co-Planar Stereotactic Atlas of the Human Brain*. New York: Thieme Medical Publishers.

Xu, P., Peng, S., Luo, Y. J., and Gong, G. (2021). Facial expression recognition: a meta-analytic review of theoretical models and neuroimaging evidence. *Neurosci. Biobehav. Rev.* 127, 820–836. doi: 10.1016/j.neubiorev.2021.05.023

frontiers | Frontiers in Behavioral Neuroscience

TYPE Opinion
PUBLISHED 04 November 2022
DOI 10.3389/fnbeh.2022.943896

Check for updates

OPEN ACCESS

EDITED BY
Rüdiger J. Seitz,
Heinrich Heine University of Düsseldorf, Germany

REVIEWED BY
Masumi Wakita,
Teikyo University, Japan

*CORRESPONDENCE
Sabine Bergner
sabine.bergner@uni-graz.at

SPECIALTY SECTION
This article was submitted to Emotion Regulation and Processing, a section of the journal Frontiers in Behavioral Neuroscience

RECEIVED 14 May 2022
ACCEPTED 03 August 2022
PUBLISHED 04 November 2022

CITATION
Bergner S, Rybnicek R and Koschutnig K (2022) Leadership and credition: Followers' neural response to leaders who are perceived as transformational.
Front. Behav. Neurosci. 16:943896.
doi: 10.3389/fnbeh.2022.943896

COPYRIGHT
© 2022 Bergner, Rybnicek and Koschutnig. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Leadership and credition: Followers' neural response to leaders who are perceived as transformational

Sabine Bergner[1]*, Robert Rybnicek[2] and Karl Koschutnig[1]

[1]Department of Psychology, University of Graz, Graz, Austria, [2]Department of Corporate Leadership and Entrepreneurship, University of Graz, Graz, Austria

KEYWORDS

credition, implicit leadership theories (ILTs), neuroimaging (anatomic and functional), transformational leadership (TL), neuroleadership

Introduction

Transformational leadership (TL) has gained much attention in current leadership research (Zhao and Li, 2019) as it results in superior organizational, team and individual performance (Wang et al., 2011). According to the follower-centric leadership approach, a leader's level of transformational behavior is not only dependent on the leader's *action* but also on the follower's *perception* and *belief* (Brown, 2018). For instance, followers who believe their leader to be more transformational—irrespective of the leader's actual behavior—show higher commitment and extra effort at work (Felfe and Schyns, 2010). This suggests that TL is also in the eye of the beholder and affected by follower beliefs (Howell and Shamir, 2005).

This fMRI-study is the first to investigate the followers' neural reaction to *perceived* transformational leadership and provides novel insights into the question why TL matters. It examines the neural patterns that are activated when followers believe a leader to be transformational and examines whether these patterns relate to the level of perceived TL. Furthermore, it investigates whether followers' neural activations predict their motivation at work.

Transformational leadership and its perception

TL describes a leadership approach that focuses on transcendent and superior goals (Antonakis and Day, 2018). At its core, it creates positive change and *transforms* followers so that they "transcend their own self-interests for the good of the group, organization, or society" (Bass, 1990, p. 53), resulting in followers doing "more than they intended and [...] even thought possible" (Bass, 1998, p. 4). To induce the intended follower transformation, leaders ought to *create an attractive future vision* (inspirational motivation), *support followers* (individualized consideration), *set high ethical standards* (idealized influence) and *stimulate followers' creative thinking* (intellectual stimulation).

Previous research has illustrated several positive follower reactions to perceiving TL. For instance, followers feel more valued and optimistic, experience more positive emotions and sense higher moral values. They feel positively connected to their leader,

experience more fairness and regard their work as more important (Pillai et al., 1999; Dirks and Ferrin, 2002; Turner et al., 2002; Bono and Judge, 2003; Kark et al., 2003; Keller, 2006; Bono et al., 2007; Tims et al., 2011). Importantly for this study, the *affect* tied to the followers' positive reaction when perceiving TL is seen as the most proximate follower reaction to TL (Ng, 2017) and might be a central reason why transformational leaders show impact.

Followers' neural reaction to transformational leadership

So far, research on followers' *neural reaction* to TL is theoretical in nature. However, neuroimaging research conducted by Schjoedt et al. (2011) and Molenberghs et al. (2017) provides initial support for the assumption that leadership, in a broader sense, activates distinct neural patterns. Even though this research focused neither on TL nor on the business context, we build on it and assume that perceiving TL triggers distinct neural activations. More detailed, we expect TL to trigger the followers' *dopaminergic reward circuit*. This assumption is based on the following considerations: First, TL provokes reactions that represent well-known affective phenomena studied in social and affective neuroscience, e.g., TL triggers follower optimism, trust, generosity and fairness, all phenomena examined in neuroscience (Davidson et al., 2009); Second, according to results from social and affective neuroscience, these phenomena trigger the dopaminergic reward circuit, e.g., individuals who perceive trust, fairness and generosity activate the ventral striatum (Mobbs et al., 2009; Izuma et al., 2010; Shenhav and Greene, 2010). Additionally, those who feel optimistic display activations in the amygdala and rostral anterior cingulate cortex (Sharot et al., 2007); Third, among the diverse mechanisms underlying TL, its rewarding value is particularly important. Those who experience TL feel rewarded (Tee, 2015), a feeling which again corresponds to activation in the dopaminergic reward circuit (Liu et al., 2011). Concluding, perceiving TL is thought to trigger the dopaminergic reward circuit.

Research focus

Based on the fact that those affective reactions reported by followers who perceive TL trigger the *dopaminergic reward circuit*, we expect that perceiving a leader to be transformational triggers the same circuit. Furthermore, we assume a positive relation between the intensity of the neural activations and the perceived level of TL, as the reward circuit activations correlate with the level of positive affect, emotion and mood (Haber and Knutson, 2010).

Hypothesis 1: Followers who believe their leader to be transformational—irrespective of the actual behavior—display activations in their dopaminergic reward circuit.

Hypothesis 2: The more followers believe their leader to be transformational, the stronger will be their neural activations.

As this is—to the best of our knowledge—the first neural study on followers' perception of TL, we also address the question of whether neuroimaging insights predict follower outcomes (Waldman et al., 2017). Therefore, we exploratively study whether the followers' neural response to perceived TL relates to their motivation at work, an outcome frequently examined in the business context.

Research question: Does the followers' neural response to perceived TL relate to their motivation at work?

Method

Forty-four (29♀, M_{age} = 25.00, SD_{age} = 2.26) healthy MBA students participated in the study. They were screened for exclusion criteria (metal implants, physical impairment, pregnancy, psychosis), provided written informed consent and received a fixed compensation (€15).

In the experiment's *pre-scanning part*, participants were told a cover story to help them establish the follower role. Accordingly, they had the chance for an internship supervised by recognized leaders (both male). Depending on their task performance in the MR-scanner they would be recruited by one of two leaders—one transformational (TL), the other not (nonTL). The better their task performance in the MR-scanner, the higher their chance for being selected by the transformational leader; contrariwise the chance for the nonTL leader arose.

Both leaders and their leadership behavior were introduced using audio vignettes and portrait pictures. Participants listened to a speech that was given by each leader and saw a portrait picture of each. Importantly, study participants did not know that (a) both leaders were fictional characters (pictures obtained from Neutralized Faces Database; Ebner, 2008), (b) the speeches were derived from Kirkpatrick and Locke's (1996) vignettes on TL/nonTL and that (c) professional announcers recorded the speeches. Leaders, speeches, announcers and portrait pictures were counterbalanced and randomized.

In the *scanning part*, an event-related design with a leadership and control treatment was conducted[1]. In every treatment, participants completed 50 of the trials depicted in Figure 1A. Each trial began with the performance task in which two circles with dots were displayed. Participants had to decide which of the circles held more dots (see Dehaene et al., 2005;

1 This study was part of a multiple-study design with various treatments (see Rybnicek et al., 2019). Here we focus only on the leadership and control treatment.

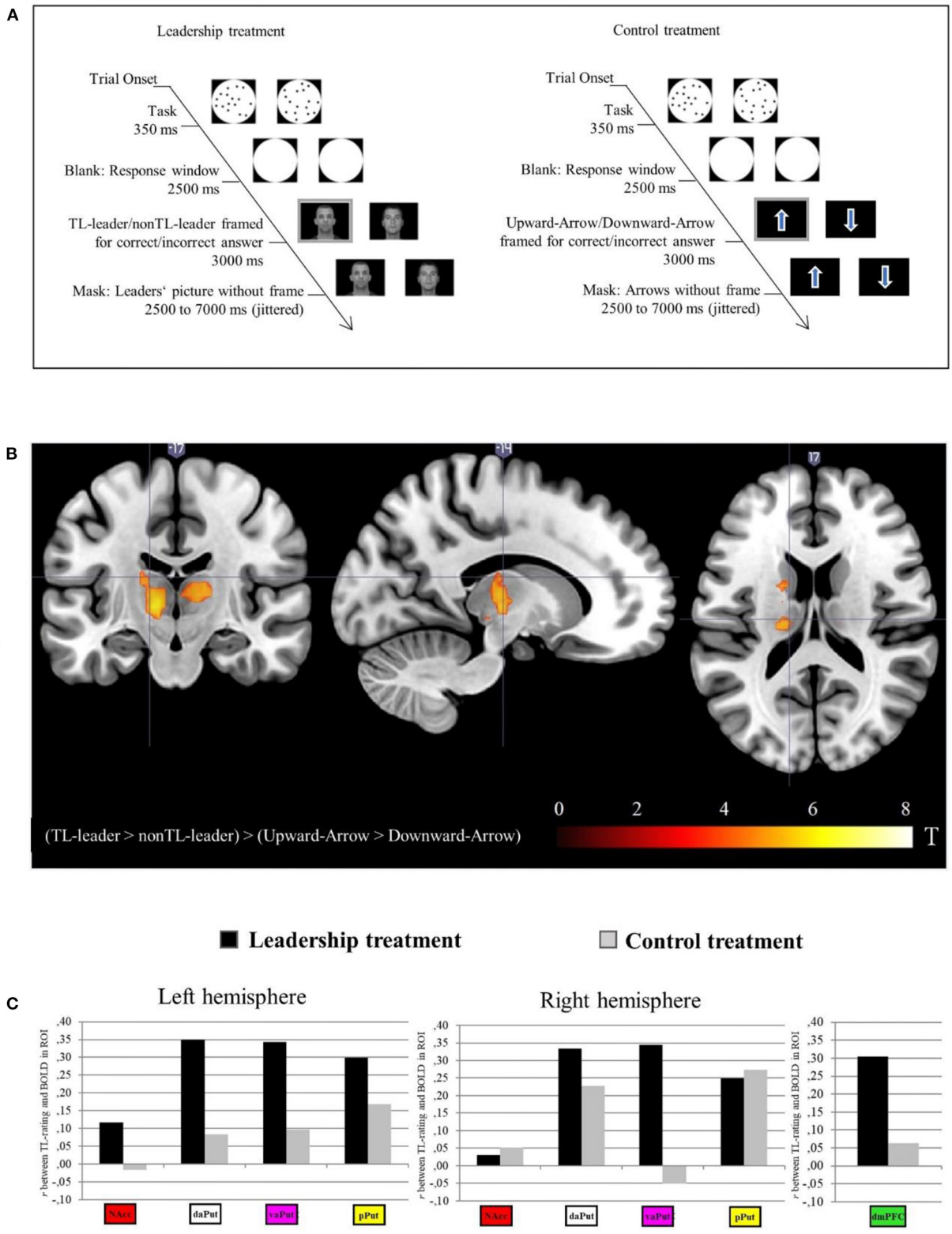


FIGURE 1
Paradigm **(A)**, whole-brain analysis **(B)** and correlation between BOLD-signal and behavioral ratings on transformational leadership **(C)**. NAcc, Nucleus accumbens; daPut, dorsal anterior putamen; vCaud, ventral caudate; vaPut, ventral anterior putamen; dCaud; dorsal caudate; pPut, posterior putamen; dmPFC, dorsomedial prefrontal cortex. Facial images reproduced with permission from the Max Planck Institute for Human Development, Center for Lifespan Psychology, Berlin, Germany, available at https://faces.mpdl.mpg.de/.

Costa et al., 2011). If the task was solved correctly, then—in the leadership treatment—the portrait picture of the TL-leader was framed, otherwise the opposing picture was framed. In the control treatment an upward-/downward-facing arrow was framed when the task was solved correctly/incorrectly. For motivational reasons, we adjusted the task so that all participants completed 60% of the trials correctly.

In the *post-scanning part*, participants rated the two leaders' TL behavior using the Multifactor-Leadership-Questionnaire (Bass and Avolio, 1995). The likability of and motivation to work for the leaders were each assessed with a single-item scale (5-point rating).

Neural activity was measured using the blood-oxygen-level-dependent (BOLD) signal. This works by detecting the changes in blood oxygenation and blood flow that occur in response to neural activity. Before the BOLD-signal was analyzed, systematic non-task-related sources of variability were removed (e.g., artifacts due to head movement). Then, general linear modeling (GLM) was used for the first-level analyses to identify an increase/decrease of the BOLD-signal in response to the treatment or baseline signal (Dimoka, 2012; Dulebohn et al., 2016). In a third step, second-level analyses were conducted to make inferences about the whole participant group. Finally, region of interest (ROI) analysis was used to focus on the activations in predefined brain areas that are central to the reward circuitry (Kätsyri et al., 2012; see Figure 1C). Analyses were corrected for multiple comparisons. Whole brain activations were family-wise-error (FWE) corrected using a voxel-level FWE of $p < 0.05$ as a measure of significance. Additionally, mean percent signal change was extracted for each ROI using MarsBaR software (Brett et al., 2002). Additional information on the MRI procedure, data acquisition, data analysis and on the method in general is provided in the Online supplement.

Results

Confirming the different TL-levels in the leader treatment, the TL-leader received higher TL-ratings than the nonTL-leader ($t_{(1,42)} = 32.55$, $p < 0.01$; $M_{\text{TL-leader}} = 4.40$, $SD_{\text{TL-leader}} = 0.34$ vs. $M_{\text{nonTL-leader}} = 1.52$, $SD_{\text{nonTL-leader}} = 0.39$). Regarding hypothesis 1, the simple activation contrast of the leadership (TL-Leader > nonTL-Leader) and control contrast (Upward-Arrow > Downward-Arrow) were studied to account for activation from answering the task correctly. Table 1 and Figure 1B reveal that followers who believe their leader to be transformational activate the putamen, thalamus and supplementary motor area, which largely supports hypothesis 1.

Regarding hypothesis 2, the BOLD response beta values for the TL-leader > nonTL-leader contrast were extracted for the predefined ROIs and correlated with the behavioral TL-ratings (MLQ-rating). Figure 1C demonstrates positive correlations between activations of the daPut (left/right), vaPut (left/right), and dCaud and TL-ratings. Importantly, the beta values of the control contrast (Upward-Arrow > Downward-Arrow) did not correlate with the TL-ratings. Thus, hypothesis 2 is largely supported.

Findings on the research question demonstrate that the followers' motivation to work for a leader significantly relates to the parameter estimates of the TL-leader > nonTL-leader in these ROIs: daPut (left $r = 0.31$, $p < 0.05$; right $r = 0.33$, $p < 0.05$), vaPut (left $r = 0.31$, $p < 0.05$; right $r = 0.33$, $p < 0.05$) and pPut (left $r = 0.31$, $p < 0.05$; right $r = 0.39$, $p < 0.05$). The activation in these ROIs explained $R^2_{\text{adj}} = 15\%$ of the variance in follower motivation ($F_{2,41} = 8.25$, $p < 0.01$). Hierarchical regressions on follower motivation also showed that the BOLD-signal for the pPut (right) added validity over behavioral TL-ratings ($\Delta R^2 = 0.07$, $p < 0.05$; controlled for leader likability). Thus, followers' neural response to perceived TL correlates with their motivation and adds incremental validity over TL-ratings when predicting motivation.

Discussion, conclusion and limitation

This study has two central findings. First, it reveals that the *pure belief* of a leader being transformational triggers distinct neural activations in the followers' reward circuitry. Second, it demonstrates that the neural response to perceived TL not only correlates with the followers' level of motivation but even predicts it beyond well-established rating measurements.

Regarding the first finding, this study revealed that followers who believed their leader to be transformational show activation in parts of their reward circuitry, which included the putamen, thalamus and SMA. These brain areas became even more strongly activated the more followers believed their leader to be transformational. Notably, neither personal interaction with nor actual behavior from the leader was necessary to trigger these brain areas. Therefore, the finding supports the *social construction perspective* of leadership (Keller, 2006), according to which leadership is partly constructed in the mind of followers and therefore to a certain extent independent of the actual leader behavior or leader–follower interaction.

Our results highlight the relevance of the reward circuitry when processing perceived TL. This is a novel insight and adds to findings from Schjoedt et al. (2011) and Molenberghs et al. (2017) who conducted the only existing fMRI-studies in the field but examined leadership in a rather general sense and neither focused on TL or the business context. As the reward circuitry is triggered when individuals experience rewarding or hedonistic values, it might be argued that followers feel rewarded or valued when processing TL. This assumption is supported by results showing that TL resembles an *idealized* leadership prototype which is loaded with appealing, rewarding

TABLE 1 *T*-values for significantly activated voxels, MNI coordinates, and cluster sizes.

Experimental effect	MNI coordinate			Voxels	Peak *T*
Hemisphere/Region	*x*	*y*	*Z*		
Main effect leadership treatment (TL-leader > nonTL-leader)					
R Caudate nucleus	18	8	−11	798	7.70
L Putamen	−21	2	−14	247	7.94
L Medial orbital frontal gyrus	3	35	−14	134	7.55
R Middle occipital gyrus	18	−103	−5	74	6.18
R Middle cingulum	0	−37	37	127	6.52
L Superior frontal gyrus	−21	32	52	41	5.90
L Cerebellum	−15	−79	−17	101	6.42
R Cerebellum	42	−70	−35	86	6.74
Main effect control treatment (upward-arrow > downward-arrow)					
R Caudate nucleus	9	8	−8	23	6.27
Contrast of the simple contrasts from the leadership and control treatment (TL-leader > nonTL-leader) > (upward-arrow > downward-arrow)					
L Putamen	−21	8	7	68	4.43
L Thalamus	−12	−16	1	103	5.35
R Thalamus	15	−19	10	62	4.63
L Supplementary motor area	−6	14	46	172	5.30

Data were corrected for multiple comparisons on a voxel-level (FWE, $p < 0.05$).

and attractive ideas about how leaders behave (Hartog et al., 1999). As such beliefs represent so-called implicit leadership theories (Eden and Leviatan, 1975), our findings not only add a neural layer to the idea that TL represents idealized leadership but also offer insights into the neural underpinning of implicit leadership theories. Additionally, the relevance of the reward circuitry neurologically supports the well-known, yet only behaviorally examined link between TL and positive follower affect like optimism, trust or generosity (e.g., Bono and Judge, 2003; Bono et al., 2007; Bregenzer et al., 2019) as these phenomena commonly trigger the dopaminergic reward circuit (Mobbs et al., 2009; Izuma et al., 2010; Shenhav and Greene, 2010).

Regarding the second finding of interest, our results demonstrate that followers' neural responses to TL correlate with their level of motivation and even predict it beyond traditional leadership ratings. Given that this study examined the neural foundation of followers' *subjective beliefs* in a leader's TL-level, this finding highlights the relevance of *beliefs* in leadership. Therefore, it also strengthens theoretical considerations of the credition model (Seitz et al., 2018), according to which belief structures shape actions and influence motivation. While existing research supports this notion—e.g., in the educational (Mitropoulou et al., 2018) and health settings (Meissner, 2017), this study primarily validates the credition model in the business context and further offers a neural underpinning thereof.

As with any study, there are limitations. As we only investigated TL, no conclusions regarding other leadership behaviors can be drawn. Furthermore, only male leaders were examined. Therefore, it remains unclear whether female leaders would trigger similar findings. Finally, no individual differences among followers were considered. As such differences affect the perception of TL (Felfe and Schyns, 2006), future studies need to elaborate the impact of these differences. Despite these limitations we feel that our findings offer an important step toward understanding the neural mechanisms underlying leadership powers.

Author contributions

SB was the PI and had the lead in writing the paper. KK did the analyses and co-wrote the paper. RR set up the experiment and co-wrote the paper. All authors contributed to the article and approved the submitted version.

Funding

This paper is funded by Rüdiger Seitz, via the Volkswagen Foundation, Siemens Healthineers, and the Betz Foundation. The external funders were not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships

that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnbeh.2022.943896/full#supplementary-material

References

Antonakis, J., Day, D. V., eds. (2018). *The Nature of Leadership*. Los Angeles, CA: SAGE Publications Inc. doi: 10.4135/9781506395029

Bass, B. M. (1990). *Bass & Stogdill's Handbook of Leadership. Theory, Research, and Managerial Applications*. New York, NY: The Free Press.

Bass, B. M. (1998). *Transformational Leadership. Industrial, Military, and Educational Impact*. New Jersey, NJ: Lawrence Erlbaum Associates.

Bass, B. M., and Avolio, B. J. (1995). *MLQ Multifactor Leadership Questionnaire. Technical Rep*ort. Redwood City, CA: Mind Garden. doi: 10.1037/t03624-000

Bono, J. E., Foldes, H. J., Vinson, G., and Muros, J. P. (2007). Workplace emotions. The role of supervision and leadership. *J. Appl. Psychol.* 92, 1357–1367. doi: 10.1037/0021-9010.92.5.1357

Bono, J. E., and Judge, T. A. (2003). Self-concordance at work. Toward understanding the motivational effects of transformational leaders. *Acad. Manag. J.* 46, 554–571. doi: 10.5465/30040649

Bregenzer, A., Felfe, J., Bergner, S., and Jiménez, P. (2019). How followers' emotional stability and cultural value orientations moderate the impact of health-promoting leadership and abusive supervision on health-related resources. *German J Human Res Manag.* 33, 307–336. doi: 10.1177/2397002218823300

Brett, M., Anton, J.-L., Valabregue, R., and Poline, J.-B. (2002). "Region of interest analysis using an SPM toolbox, "in: *Presented at the 8th International Conference on Functional Mapping of Human Brain Sendai, Japan, June 2002* (Sendai).

Brown, D. J. (2018). "In the minds of followers. Follower-centric approaches to leadership," in: *The Nature of Leadership, eds.* J. Antonakis, D. V. Day (Los Angeles, CA: SAGE Publications Inc), 82–108. doi: 10.4135/9781506395029.n4

Costa, A. J., Silva, J. B. L., Chagas, P. P., Krinzinger, H., Lonneman, J., Willmes, K., et al. (2011). A hand full of numbers. A role for offloading in arithmetics learning? *Front. Psychol.* 2, 368–372. doi: 10.3389/fpsyg.2011.00368

Davidson, R. J., Sherer, K. R., and Goldsmith, H. H. (2009). *Handbook of Affective Sciences*. New York, NY: Oxford University Press.

Dehaene, S., Izard, V., and Piazza, M. (2005). *Control over non-numerical parameters in numerosity experiments*. Available online at: www.unicog.org/docs/DocumentationDotsGeneration.doc (accessed July 25, 2022).

Dimoka, A. (2012). How to conduct a functional magnetic resonance (fMRI) study in social science research. *MIS Q.* 36, 811–840. doi: 10.2307/41703482

Dirks, K. T., and Ferrin, D. L. (2002). Trust in leadership. Meta-analytic findings and implications for research and practice. *J. Appl. Psychol.* 87, 611–628. doi: 10.1037/0021-9010.87.4.611

Dulebohn, J. H., Davison, R. B., Lee, S. A., Conlon, D. E., McNamara, G., and Sarinopoulos, I. C. (2016). Gender differences in justice evaluations. Evidence from fMRI. *J. Appl. Psychol.* 101, 151–170. doi: 10.1037/apl0000048

Ebner, N. C. (2008). Age of face matters. Age-group differences in ratings of young and old faces. *Behav. Res. Methods* 40, 130–136. doi: 10.3758/BRM.40.1.130

Eden, D., and Leviatan, U. (1975). Implicit leadership theory as a determinant of the factor structure underlying supervisory behavior scales. *J. Appl. Psychol.* 60, 736–741. doi: 10.1037/0021-9010.60.6.736

Felfe, J., and Schyns, B. (2006). Personality and the perception of transformational leadership. The impact of extraversion, neuroticism, personal need for structure, and occupational self-efficacy. *J. Appl. Soc. Psychol.* 36, 708–739. doi: 10.1111/j.0021-9029.2006.00026.x

Felfe, J., and Schyns, B. (2010). Followers' personality and the perception of transformational leadership. Further evidence for the similarity hypothesis. *Br. J. Manag.* 21, 393–410. doi: 10.1111/j.1467-8551.2009.00649.x

Haber, S. N., and Knutson, B. (2010). The reward circuit. Linking primate anatomy and human imaging. *Neuropsychopharmacology* 35, 4–26. doi: 10.1038/npp.2009.129

Hartog, D. N., House, R. J., Hanges, P. J., Ruiz-Quintanilla, S., Dorfman, P. W., et al. (1999). Culture specific and cross-culturally generalizable implicit leadership theories. Are attributes of charismatic/transformational leadership universally endorsed? *Leadership Q.* 10, 219–256. doi: 10.1016/S1048-9843(99)00018-1

Howell, J. M., and Shamir, B. (2005). The role of followers in the charismatic leadership process. Relationships and their consequences. *Acad. Manag. Rev.* 30, 96–112. doi: 10.5465/amr.2005.15281435

Izuma, K., Saito, D. N., and Sadato, N. (2010). Processing of the incentive for social approval in the ventral striatum during charitable donation. *J. Cogn. Neurosci.* 22, 621–631. doi: 10.1162/jocn.2009.21228

Kark, R., Shamir, B., and Chen, G. (2003). The two faces of transformational leadership. Empowerment and dependency. *J. Appl. Psychol.* 88, 246–255. doi: 10.1037/0021-9010.88.2.246

Kätsyri, J., Hari, R., Ravaja, N., and Nummenmaa, L. (2012). The opponent matters. Elevated fMRI reward responses to winning against a human versus a computer opponent during interactive video game playing. *Cereb. Cortex* 23, 2829–2839. doi: 10.1093/cercor/bhs259

Keller, R. T. (2006). Transformational leadership, initiating structure, and substitutes for leadership. A longitudinal study of research and development project team performance. *J. Appl. Psychol.* 91, 202–210. doi: 10.1037/0021-9010.91.1.202

Kirkpatrick, S. A., and Locke, E. A. (1996). Direct and indirect effects of three core charismatic leadership components on performance and attitudes. *J. Appl. Psychol.* 81, 36–51. doi: 10.1037/0021-9010.81.1.36

Liu, X., Hairston, J., Schrier, M., and Fan, J. (2011). Common and distinct networks underlying reward valence and processing stages. A meta-analysis of functional neuroimaging studies. *Neurosci. Biobehav. Rev.* 35, 1219–1236. doi: 10.1016/j.neubiorev.2010.12.012

Meissner, K. (2017). "Believing in the effectiveness of treatment. From placebo to credition and back," in: *Processes of Believing. The Acquisition, Maintenance, and Change in Creditions,* eds. H.-F. Angel, L. Oviedo, R. F. Paloutzian, A. Runehov, R. J. Seitz (Cham: Springer International Publishing), 125–137. doi: 10.1007/978-3-319-50924-2_9

Mitropoulou, V., Gkirlou, D., and Meke, M. (2018). Building worldviews with credition teaching model in high school curricula of religious education. A controlled experimental study. *Soc. Study* 8, 325–335. doi: 10.17265/2159-5526/2018.07.003

Mobbs, D., Yu, R., Meyer, M., Passamonti, L., Seymour, B., Calder, A. J., et al. (2009). A key role for similarity in vicarious reward. *Science* 324, 900. doi: 10.1126/science.1170539

Molenberghs, P., Prochilo, G., Steffens, N. K., Zacher, H., and Haslam, S. A. (2017). The neuroscience of inspirational leadership. The importance of collective-oriented language and shared group membership. *J. Manage.* 43, 2168–2194. doi: 10.1177/0149206314565242

Ng, T. W. (2017). Transformational leadership and performance outcomes. Analyses of multiple mediation pathways. *Leadership Q.* 28, 385–417. doi: 10.1016/j.leaqua.2016.11.008

Pillai, R., Schriesheim, C. A., and Williams, E. S. (1999). Fairness perceptions and trust as mediators for transformational and transactional leadership. A two-sample study. *J. Manage.* 25, 897–933. doi: 10.1177/014920639902500606

Rybnicek, R., Bergner, S., and Gutschelhofer, A. (2019). How individual needs influence motivation effects: a neuroscientific study on McClelland's need theory. *Rev. Manag. Sci.* 13, 443–482. doi: 10.1007/s11846-017-0252-1

Schjoedt, U., Stødkilde-Jørgensen, H., Geertz, A. W., Lund, T. E., and Roepstorff, A. (2011). The power of charisma. Perceived charisma inhibits the frontal executive network of believers in intercessory prayer. *Soc. Cogn. Affect. Neurosci.* 6, 119–127. doi: 10.1093/scan/nsq023

Seitz, R. J., Paloutzian, R. F., and Angel, H.-F. (2018). From believing to belief. A general theoretical model. *J. Cogn. Neurosci.* 30, 1254–1264. doi: 10.1162/jocn_a_01292

Sharot, T., Riccardi, A. M., Raio, C. M., and Phelps, E. A. (2007). Neural mechanisms mediating optimism bias. *Nature* 450, 102–105. doi: 10.1038/nature06280

Shenhav, A., and Greene, J. D. (2010). Moral judgments recruit domain-general valuation mechanisms to integrate representations of probability and magnitude. *Neuron* 67, 667–677. doi: 10.1016/j.neuron.2010.07.020

Tee, E. Y. (2015). The emotional link. Leadership and the role of implicit and explicit emotional contagion processes across multiple organizational levels. *Leadership Q.* 26, 654–670. doi: 10.1016/j.leaqua.2015.05.009

Tims, M., Bakker, A. B., and Xanthopoulou, D. (2011). Do transformational leaders enhance their followers' daily work engagement? *Leadership Q.* 22, 121–131. doi: 10.1016/j.leaqua.2010.12.011

Turner, N., Barling, J., Epitropaki, O., Butcher, V., and Milner, C. (2002). Transformational leadership and moral reasoning. *J. Appl. Psychol.* 87, 304–311. doi: 10.1037/0021-9010.87.2.304

Waldman, D. A., Ward, M. K., and Becker, W. J. (2017). Neuroscience in organizational behavior. *Annu. Rev. Organ. Psychol. Organ. Behav.* 4, 425–444. doi: 10.1146/annurev-orgpsych-032516-113316

Wang, G., Oh, I.-S., Courtright, S. H., and Colbert, A. E. (2011). Transformational leadership and performance across criteria and levels. A meta-analytic review of 25 years of research. *Group Organ. Manag.* 36, 223–270. doi: 10.1177/1059601111401017

Zhao, H., and Li, C. (2019). A computerized approach to understanding leadership research. *Leadership Q.* 30, 396–416. doi: 10.1016/j.leaqua.2019.06.001

frontiers | Frontiers in Behavioral Neuroscience

TYPE Systematic Review
PUBLISHED 17 October 2022
DOI 10.3389/fnbeh.2022.928522

Check for updates

OPEN ACCESS

EDITED BY
Rüdiger J. Seitz,
Heinrich Heine University of Düsseldorf, Germany

REVIEWED BY
Ferdinand Binkofski,
RWTH Aachen University, Germany
Elke R. Gizewski,
Innsbruck Medical University, Austria

*CORRESPONDENCE
Bruno J. Weder
bj.weder@bluewin.ch

SPECIALTY SECTION
This article was submitted to Emotion Regulation and Processing, a section of the journal Frontiers in Behavioral Neuroscience

RECEIVED 25 April 2022
ACCEPTED 11 July 2022
PUBLISHED 17 October 2022

CITATION
Weder BJ (2022) Mindfulness in the focus of the neurosciences - The contribution of neuroimaging to the understanding of mindfulness. *Front. Behav. Neurosci.* 16:928522. doi: 10.3389/fnbeh.2022.928522

COPYRIGHT
© 2022 Weder. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Mindfulness in the focus of the neurosciences - The contribution of neuroimaging to the understanding of mindfulness

Bruno J. Weder*

Support Centre for Advanced Neuroimaging (SCAN), Institute for Diagnostic and Interventional Neuroradiology, Inselspital, Bern University Hospital, University of Bern, Bern, Switzerland

Background: Mindfulness affects human levels of experience by facilitating the immediate and impartial perception of phenomena, including sensory stimulation, emotions, and thoughts. Mindfulness is now a focus of neuroimaging, since technical and methodological developments in magnetic resonance imaging have made it possible to observe subjects performing mindfulness tasks.

Objective: We set out to describe the association between mental processes and characteristics of mindfulness, including their specific cerebral patterns, as shown in structural and functional neuroimaging studies.

Methods: We searched the MEDLINE databank of references and abstracts on life sciences and biomedical topics *via* PubMed using the keywords: "mindfulness," "focused attention (FA)," "open monitoring (OM)," "mind wandering," "emotional regulation," "magnetic resonance imaging (MRI)" and "default mode network (DMN)." This review extracted phenomenological experiences across populations with varying degrees of mindfulness training and correlated these experiences with structural and functional neuroimaging patterns. Our goal was to describe how mindful behavior was processed by the constituents of the default mode network during specific tasks.

Results and conclusions: Depending on the research paradigm employed to explore mindfulness, investigations of function that used fMRI exhibited distinct activation patterns and functional connectivities. Basic to mindfulness is a long-term process of learning to use meditation techniques. Meditators progress from voluntary control of emotions and subjective preferences to emotional regulation and impartial awareness of phenomena. As their ability to monitor perception and behavior, a metacognitive skill, improves, mindfulness increases self-specifying thoughts governed by the experiential phenomenological self and reduces self-relational thoughts of the narrative self. The degree of mindfulness (ratio of self-specifying to self-relational thoughts) may affect other mental processes, e.g., awareness, working memory, mind wandering and belief formation. Mindfulness prevents habituation and the constant assumptions associated with mindlessness. Self-specifying thinking during mindfulness and self-relational thinking in the narrative self relies on the default mode network. The main constituents of this network are the dorsal and medial prefrontal cortex, and posterior

cingulate cortex. These midline structures are antagonistic to self-specifying and self-relational processes, since the predominant process determines their differential involvement. Functional and brain volume changes indicate brain plasticity, mediated by mental training over the long-term.

KEYWORDS

mindfulness, mind wandering, focused attention, open monitoring, self-specifying processes, magnetic resonance imaging, default mode network

Introduction

Through mindfulness, we discern new aspects of experience, rather than seeking to confirm established convictions when we assimilate our experience (Langer and Moldoveanu, 2000). In mindfulness, "what" is not as important as "how:" immediate experience is paramount. Active recognition of the new sensations we perceive binds our attention to the present, increasing our awareness of the context of our activities and our perspectives on them. The work of freeing ourselves from the models and categories of the past brings a new awareness: a subjective feeling of involvement in ongoing events and more intense experience of the "Here and Now."

Since the early 1970s, mindfulness studies have elucidated the basic characteristics of mindfulness (Brown et al., 2007). These are (1) dedication to immediate experience and concentration on the present, (2) assuming a dispassionate attitude that allows instantaneous assessment of what is observed, and (3) appreciation and acceptance of sensations, feelings, or thoughts as they arise. This fundamentally impartial attitude allows individuals to anticipate intrusion of obstructive rumination and subjective values or preferences, increasing the flexibility of their thought and improving coping strategies in overwhelming or stressful situations and strengthening resilience (Keye and Pidgeon, 2013).

Abbreviations: FA, Focused Attention; OM, Open monitoring; SIT, Stimulus Independent Thought; SOT, Stimulus Oriented Thought; MW, Mind Wandering; EES, Enactive Experiential Self; EPS, Experiential phenomenological Self; NS, Narrative Self; MRI, Magnetic Resonance Imaging; s-MRI, structural Magnetic Resonance Imaging; act-fMRI, functional Magnetic Resonance Imaging during activation task; rs-fMRI, functional Magnetic Resonance Imaging during resting state; BOLD, Blood Oxygen Level Dependent; DMN, Default Mode Network; FPCN, Fronto-parietal control network; mPFC, medial Prefrontal Cortex; lPFC, lateral Prefrontal Cortex; PCC, Posterior Cingulate Cortex; ACC, Anterior Cingulate Cortex; AIC, Anterior Insular Cortex; OFC, Orbitofrontal Cortex; IPL, Inferior Parietal Lobule; RSP Cortex, Retro-Splenial Cortex; PHG, Para-Hippocampal Gyrus; HIC, Hippocampus; TPJ, Temporo-parietal Junction; KIMS, Kentucky Inventory of Mindfulness Skills; MAAS, Mindful attention Awareness Scale; FMI, Freiburg Mindfulness Inventory; FFMQ, Five Facet Mindfulness Questionnaire.

The first studies often sought to determine the effects of mindfulness on health. Kabat-Zinn et al. (1985) investigated the significance of mindfulness to the self-control of pain. Shapiro and Schwartz (2000) proposed that mindfulness helped reduce stress [Mindfulness Based Stress Reduction, MBSR] and developed an integrated model for stress reduction and health. Teasdale et al. (2000) thought mindfulness opened the possibility one could monitor one's own cognitive processes and could aid in treating mental illnesses. They found the relapse rate for depression decreased when patients were treated with behavioral therapy and mindfulness (Teasdale et al., 2002).

Industry had a separate early interest: applying techniques to motivate workers and managers to practice mindfulness. Industry studies of mindfulness showed the practice fostered creativity and helped reduce burnout (Goodman and Schorling, 2012; Langer, 2014). Davenport and Pagnini (2016) elucidated the conflict between mindfulness and indifference in education with the goal of promoting more mindful learning in the schools.

The scientific coordinator and moderator of the first Mind-and-Life-Dialogs[1] (Varela, 1996) proposed that the meditative potential of human experience is a necessary complement to the inner representations of the external world posited by cognitive sciences. He conceived of neurophenomenology as a reciprocal relationship between phenomenological access and experiential structures and inner representations of the external world. Varela et al. (2016) asked this neurophenomenological question: Can classical cognitive science assess the experiential content of mental states in the philosophical term qualia? He saw the meditation of Buddhist monks, characterized by their ability to be present in body and mind, as the exemplar of a rigorous paradigm for investigating experience at first hand. In the Buddhist tradition, meditation is a concerted act of body and mind, which receive perceptions while the individual's conscious awareness remains directed, uninterrupted, at the object perceived (Revel and Ricard, 1999). Buddhist practice is also confident of the stream of consciousness (James, 1890) described as having five characteristics: (1) subjectivity; (2) permanent change; (3)

1 Dharamsala.

continuity; (4) autonomy regarding the specific mental activity; and (5) preference for distinct objects.

About 20 years ago, separate research groups identified brain areas, mainly at midline cortices, in which BOLD (blood oxygen level dependent) signals decreased as BOLD activity in areas involved in goal-directed behaviors increased (Shulman et al., 1997; Gusnard and Raichle, 2001; Mazoyer et al., 2001). Gusnard and Raichle (2001) recognized that these areas, which constitute the default mode network (DMN), are active at baseline and become more active during self-referential tasks. These discoveries heralded a new era and a new perspective: we could now use imaging techniques to explore mental processes like mindfulness. Garrison et al. (2015) showed that in experienced meditators DMN is typically characterized by suppression of default mode processing during meditation, beyond the resting state function observed during another active, self-relational cognitive task. This and altered baseline during rest seem to be unique features of long-term meditation.

This review is based on five themes found in ongoing conceptual and theoretical disputes on mindfulness. These themes are associated with both behavioral patterns and neuroimaging data: (1) proposals in the literature for an unequivocal operational definition of mindfulness; (2) experimental requirements derived from operational definitions met in corresponding neuroimaging projects; (3) dimensions of mindfulness and its connections to other mental processes, e.g., believing; (4) contributions of neuroimaging studies to our understanding of mindfulness, specifically studies on DMN and areas engaged in interoception and exteroception; and (5) the dynamic association between expertise in mindfulness meditation and related morphological and functional imaging patterns.

The five thematic sections of the review focus on pressing questions and observations: 1. "The quest for an operational definition of mindfulness—a semantic issue;" 2. "Dimensions of believing;" 3. "Dimensions of believing and their interrelation with mindfulness," delimiting the distinction between perceptual processing and mindfulness *via* an evaluative component; 4. "Studying mindfulness with neuroimaging using MRI—concepts and technical aspects;" 5. "Neuroimaging studies of mindfulness—shaping the brain in parallel with the experience of mindfulness meditation." The search terms we used to identify the papers we included in the review are summarized in the Supplementary material in partitions that align with these sections (Supplementary Table S1).

This review extracts phenomenological experiences across populations with varying degrees of experience in mindfulness and correlates experience with structural and functional neuroimaging patterns that reflect the way mindful behavior is processed by the constituents of the default mode network during specific tasks.

The quest for an operational definition of mindfulness—A semantic issue

The prerequisite for an operational definition of mindfulness is adherence to descriptive language that permits exploration and understanding of neuroimaging results across disciplines (Lutz et al., 2015). Bishop et al. (2004) reported proposals for an operational definition made at a consensus conference. The primary concern of attendees was specifying the essential components that would serve as the basis for verifiable predictions. The secondary concern was characterizing measures to validate the resulting construct. These efforts produced a model of mindfulness comprising two elements: (1) regulating mindfulness to heighten awareness of mental activity during an experience; (2) deciding to focus one's attention on one's immediate experience, characterized by an attitude of curiosity, openness and receptivity. The conference's initiatives aroused prompt attention and provoked ongoing discussion about developing this model. Discussants expressed reservations about the model's completeness, asking if the model comprised all the necessary components. Kabat-Zinn (1990), who defined mindfulness as paying purposeful attention characterized by immediacy and withholding judgment, proposed a different but similar two-component model.

Mindfulness creates a mental state in which one attends to sensations, feelings, and thoughts as they emerge in the stream of consciousness. The skill of observing impartially and immediately can be learned within the framework of meditation. Lutz et al. (2015) delineated four contextual features of meditation: physical posture; non-aversive affect; axiological framework; and maintenance or retention of experience. Physical postures that facilitate meditation techniques include sedentary practices like FA, OM, and ethical enhancement, and also some that require physical exercise like Hatha Yoga (Vago, 2014). The concepts behind Hatha Yoga and ethical enhancement are complex and explicitly extend mindfulness into the external world. They include empathy and demand integrity, which is particularly important in mindfulness-based research (Crane and Hecht, 2018). Ethical and religious contemplation are aspects of an axiological framework that transcend the secular perspective, so they are not a subject of this review.

There is a consensus to classify meditation techniques at least into the categories of FA meditation, OM meditation, and compassion or loving-kindness practices (Fox et al., 2016). FA-meditation focuses attention on a specific bodily act like breathing supported by a controlled posture and should sustain awareness of current experience (Bishop et al., 2004). Claims that this meditative practice predicts introspective accuracy are supported by subjective reports and objective measures of tactile sensitivity, e.g., 2D discrimination or adjusted cortical

activity (Fox et al., 2012). Mantra recitation meditation may appear similar to FA meditation but distinct through the inherent focus, i.e., the repetition of a sound, word or sentence spoken aloud or silently (Travis, 2014). OM-meditation is also introspective, characterized by curious and deliberately unrestricted receptivity to experiences, primarily physically but also mentally. The trainee gains flexibility, can regulate and sustain attention, deal immediately with experiences, and keeps an open mind (Posner, 1980). Loving-kindness meditation and compassionate meditation are closely related to each other but differ in their intention: In loving-kindness meditation subjects intend to generate sympathetic feelings for all living beings whereas in compassion meditation they cultivate empathic attitudes and behaviors to the suffering of others (Fox et al., 2016). Components of mindfulness are evident in these practices, like sustained attention and open examination of immediate experience, but they are not the same as mindfulness, which is an internalized disposition to accept experience and retain its context in daily life after meditation training (Gethin, 2011).

FA-meditation is directed to somaesthesia, since it is restricted to interoception, but OM-meditation also encompasses the external world, induced by an open-minded and curious basic attitude. The different traditions of Buddhism emphasize either awareness/mindfulness or concentration/absorption, both of which affect and promote behaviors (Mikulas, 2011). Experienced meditators perform better on the Wilkins counting test than inexperienced meditators (Wilkins et al., 1987; Valentine and Sweet, 1999). Young adults who practiced OM and FM for a week had better executive function scores on an emotional variant of the Attention Network Test than those in a control group that practiced only relaxation techniques (Ainsworth et al., 2013). Different meditation techniques can also produce differences in behavior. Colzato et al. (2012) found that subjects proficient in OM-meditation may be better at tasks that require divided attention than subjects proficient in FA-meditation. Lutz et al. (2008) suggested that FA-meditation promotes effortless concentration while OM-meditation promotes objectless attention.

Physiological effects of mindfulness were also observed, namely decreased neural response to stimuli. Brown et al. (2013) showed that after viewing stimulating images (pleasant or unpleasant), mindful intervention helped decrease amplitude of late positive potential (LPP) after 400 ms. Since LPP is an electrophysiological marker for the emotional valence of a stimulus, this decrease indicates that mindfulness modulates emotion in an early phase of generation, before cognitive suppression can inhibit explicitly expressive behavior (Gross, 2001; Sheppes and Gross, 2011). As an example of implicit sensory-affective-motor processing, it is possible the underlying enactive experiential network (Vago and Silbersweig, 2012) sustains an equanimous frame of mind. Emotion control strategies appear to change with age: people's ability to appraise positively improves with age, while their ability to suppress emotional behavior is maintained, and their ability to implement detached reappraisal declines with age (Shiota and Levenson, 2009).

The quest for an operational definition of mindfulness persisted after the consensus conference. Bishop et al. (2004)'s aim was to distinguish core elements of the model and demarcate collateral features that might actually be beneficial. Shapiro and Schwartz (2000) discussed collateral features including patience, trust, self-restraint, wisdom, and compassion. Some conjectured these benefits emerge as an individual becomes conscious of his own thoughts in a process of continuous mental cultivation (Hölzel et al., 2011; Lutz et al., 2015). Meta-awareness then facilitates the prerequisite of experiential retention. Brown et al. (2007) contended that mindfulness is an attribute of consciousness that pertains to perception rather than cognition. They assert that the distinction between mindfulness as such and the various meditative practices that help practitioners attain mindfulness are insufficient, since these practices emphasize the perception of internal phenomena. Langer (2014) suggested that the definition of mindfulness be extended from awareness of the internal world of sensory stimulation, emotions, or thoughts to attention to the outside world of salient events. This extension opens new possibilities in exercising mindfulness: creation of new categories; a proper world of thoughts freed of old patterns; openness to new experience from a first- or third-person perspective; and recognizing alternative adaptive possibilities in specific circumstances.

A core element of mindfulness is impartial registration of sensations, feelings, and thoughts (Rimes and Wingrove, 2011). Maintaining non-aversive affect is essential to mindfulness, as it may foster positive attitudes like acceptance, loving kindness, compassion, or aesthetic appreciation (Shapiro et al., 2006; Lutz et al., 2015). A responsive positive emotional state like compassion is a prerequisite for open-mindedness and curiosity because it reduces preoccupation with repetitive negative thoughts and redundant speculations (Takano and Tanno, 2009; McEvoy et al., 2010). These thoughts and speculations can afflict the subject with reiterative and critical discourse about the meaning of perceptions and disrupt attention to immediate experience. Open-mindedness, in contrast, facilitates perception during the stream of consciousness and thus sustains flexibility (Moore and Malinowski, 2009). Greenberg et al. (2010) showed that internalizing mindfulness sustained flexibility in those who solved Luchins' water jar test.

After an experience and depending on the degree of activity, the focus of attention in the daily stream of consciousness alternates between attention directed to external events and attention to internal states (James, 1890). Based on stimulus-oriented thoughts (SOT), mindful mental states subserve external perceptions. Individuals often cannot maintain SOTs and lapse into stimulus-independent thoughts (SITs), which

are associated with ruminations about past and future and their effects on self-imagination and expectations (Mason et al., 2007a,b). The boundary between mindfulness *sensu stricto* and such self-monitoring is critical. These states of self-focused attention may also cause distress (Nolen-Hoeksema, 1991; Trapnell and Campbell, 1999; Neff, 2003). Such intrusions in our consciousness, experienced as daydreams, can accompany physical activity and correlate negatively with its intensity (Killingsworth and Gilbert, 2010). Self-referential thinking and adherence to old, ingrained attitudes and patterns of thought hinders attention to outer experience and shifts the focus of internal mental activity from impartial, open perception of feelings and thoughts to the irreal realm of wishes and fears (Dambrun and Ricard, 2011). The experiential system loses its immediate relation to reality, reducing the capacity of the affected individual to adapt to new situations or conditions (Rummel and Boywitt, 2014). The lapse into an illusionary world when the mind "wanders" then pervades our thoughts. A careful, web-based study of 2,250 arbitrarily selected individuals indicated that mind wandering prevailed 46.9% of the time (Killingsworth and Gilbert, 2010). During externally directed activity, lapses into mind wandering occurred for at least 30% of the period measured.

A multilevel regression showed that individuals felt unhappier when their minds wandered. The authors concluded, "A wandering mind is an unhappy mind." The regression indicated that the content of their thoughts was a better predictor of the individual's happiness than what they were doing. Mason et al. (2007a) ascertained, in interviews with subjects at rest immediately after a functional MRI session, that mind wandering was common. Participants engaged in stimulus-oriented thinking 26% of the time, focused on their physical state 15% of the time, and were preoccupied with stimulus-independent thoughts 59% of the time. Stimulus-independent thoughts focused on the future 26% of the time, on the past 23% of the time, and were unspecified 10% of the time. By nature, mind wandering tends to inattention, contrasting clearly with the decentred, unconstrained attitude of mindfulness. Others have also reported predominantly negative effects of mind wandering, especially when the mind wanders to past events (Smallwood and O'Connor, 2011; Stawarczyk et al., 2013). One contradictory study observed that positive mood effects and mind wandering were reciprocal (Smallwood et al., 2009).

Scales for the assessment of mindfulness and inattention

Validated mindfulness scales characterize subjective dispositions or traits and describe a subject's tendency to be mindful in daily life. These scales include the Mindful attention Awareness Scale (MAAS) (Brown and Ryan, 2003), the Freiburg Mindfulness Inventory (FMI) (Walach et al., 2006), the Kentucky Inventory of Mindfulness Skills (KIMS) (Baer et al., 2004), the Five Facet Mindfulness Questionnaire (FFMQ) (Baer et al., 2006) and the Imaginal Processes Inventory (http://neuroinformatics.harvard.edu/w/public/images/5/55/Ipi.pdf). The Imaginal Processes Inventory assesses the risk an individual will lapse from mindfulness into daydreams or mind wandering (Supplementary Table S2: Scales measuring mindful traits). The Toronto mindfulness scale measures mental states; this scale can be correlated with the dispositions of mindfulness described above (Lau et al., 2006). All these are Likert scales, and all these single factor and multifactor scales contain ambiguities that must be resolved in future studies. A pressing question is how much their results depend on meditation experience and whether their findings are consistent with experimental mindfulness tasks (Bergomi et al., 2013; Chiesa, 2013). We also cannot be sure if mindful traits are related or independent factors.

Dimensions of mindfulness

A dual concept of personality theory posits that human behavior relies on two information processing systems that work in parallel and interact. These are the rational and experiential frameworks, e.g., Epstein (2003)'s cognitive—experiential self-theory (CEST). The two frameworks support first and third person perspectives. A hypothesis of classical cognitive science is that the first person perspective of empirical self-observation is a subject's privileged account of their own experience. This privileged view is inaccessible to other observers and thus irreducible to third person data (Searle, 1994). Third person perspective reduces perceptions to objects and processes that exist outside, and thus are independent of the subject's mind. Third person perspectives provide data about the objective structure and dynamics of physical systems (Chalmers, 2013). The classical assumption is dualistic, posing a dichotomy between indirectly ascertainable objects and processes and internally experienced percepts.

Scholars are increasingly criticizing the presumption of a dichotomy. Choifer (2018) partially resolves this dilemma by linguistically linking these two perspectives to the personal pronoun. He proposes that the subject exhibits two modes of consciousness: reflective or non-reflective. These two modes of being in the world allow the subject to occupy one of two mutually exclusive perspectives at a given time. In the reflective mode, the third person perspective is scientifically accessible. When the subject detaches from self-referential thoughts, they may take an experiential attitude to self-observation, the precondition for metacognitive skills like monitoring one's own perceptions (Pasquali et al., 2010).

Varela and Shear (1999) integrate the first person perspective with non-reflective thoughts as lived experience associated with cognitive and mental events into a science of consciousness. They acknowledge that this lived experience must be substantiated by third-person studies. According to this theory, the perspective of a second person (e.g., an experienced tutor) may mediate between the perspectives of the first and third person (Pauen, 2012). The discussion of the theoretical framework that supports the argument for an intermediate, second person perspective is out of scope of the review but validating second-person methods for studying human consciousness would require first establishing objective methods for comparing results across different subjects and tutors (Olivares et al., 2015). A new cognitive science construct derived from probabilistic models and based on prediction coding and the free energy principle treats both the metaphysical self ("I," the subject of experience) and the phenomenal self ("me," the object of experience) as if they occupied different levels of the phenomenal self-model (Metzinger, 2018). This distinction is purely pragmatic; it forfeits the subject of experience and asserts that an experience can be owned (Wozniak, 2018). The main questions are now: How may we describe and reliably estimate the metacognitive competence of self-monitoring? How can we grasp effects of mindfulness on the subject's experience? And how shall we integrate aspects of mindfulness vs. mindlessness into a neurobiological concept that matches patterns visible in neuroimaging that could be associated with these states?

To generate scientific hypotheses (Lutz et al., 2015) proposed a heuristic tool: orientation on a phenomenological matrix of mindfulness. They suggested we could map focused attention, open monitoring, mind wandering, and rumination within three-dimensional space by measuring object orientation, dereification, and meta-awareness. OM could be clearly differentiated from FA in experienced subjects along the axis of object orientation [1], like OM and FA can be differentiated from mind wandering along the axis of dereification [2] and OM can be differentiated from mind wandering along the axis of meta-awareness (the indication for monitoring of experience) [3]. One could also map the secondary dimensions of aperture of the focus of interest, clarity of the percept, stability of disposition and effort, each of which would reveal qualities of these mental states. Based on self-reports from a neurophenomenological experiment, OM and FA were distinct. The distinction was illustrated by the broad range of attention novices and experienced meditators exhibited during OM (Abdoun et al., 2019). FA is less clear (less vivid experience) and less stable (experiences of shorter duration) than OM, because FA relies on theoretical background and experiential knowledge to calm and slow down the mind (Revel and Ricard, 1999). Dereification is associated with impaired ability to discriminate between mental phenomena and depictions of reality. Together with unavailable meta-awareness, dereification characterizes mind-wandering, impacting negatively on wellbeing (Dahl et al., 2015).

Christoff et al. (2016) presented a complementary two-dimensional framework that places mental states subjected to deliberate constraints on one axis and mental states subjected to automatic constraints on the other. Subjects can exert cognitive control to govern their mental states deliberately (Miller, 2000). Cognitive control is most strongly exerted during goal-directed thought, less common during creative thinking and mind wandering, and least common during dreams. Automatic constraints are fundamentally different because they cannot be controlled by cognition and are most likely driven by affective and sensory salience (Todd et al., 2012), e.g., ruminations, obsessive thoughts, and addictive cravings.

Five days of integrative body-mind training grounded in traditional Chinese medicine, which included breathing adjustment and mindfulness training, gradually cultivated effortless attention and improved conflict resolution, as measured by the attention network test (ANT) (Tang et al., 2007). A study that tested the threshold for conscious perception and working memory capacity found that in meditation novices who engaged in mindfulness-based stress reduction, these capacities improved significantly more than in those who practiced alternative strategies; however, it was impossible to strictly differentiate confounding factors like test effort and stress reduction not caused by mindfulness (Jensen et al., 2012). As compared to a distinct focused attention awake state, measuring the relative concentrations of brain metabolites using ^{31}P Magnetic Resonance Spectroscopy indicated an enhanced energetic state induced by a FA meditation state in the basal ganglia and temporal lobes and, furthermore, a down-regulation of ATP-turnover in the occipital and frontal lobes after a 7 weeks training (Galijašević et al., 2021).

Long-term practice of meditation within the Tibetan Buddhist tradition cultivates a special form of attentional expertise (Brefczynski-Lewis et al., 2007) in which practitioners can sustain attention on an external or internal object over time. This is one-pointed concentration: when a state of equanimity is achieved, the dichotomy of subject and object may eventually disappear (Revel and Ricard, 1999). One study examining the effects of 3 months of systematic mental training in concentration meditation on information processing found that the practice seems to ameliorate the so-called "attentional blink deficit" in which two targets compete for limited attentional resources (Slagter et al., 2007).

Vago and Silbersweig (2012) propose a comprehensive conceptual framework to describe the functional relationship between mindfulness processing and neurobiological mechanisms: S-ART (Self-Awareness, -Regulation, and -Transcendence). Its constituents are the task positive networks (cf. attention to the external world) of the enactive experiential self (EES) and of the experiential phenomenological self (EPS), the task negative network (cf. internally directed mentation)

of the narrative self (NS), and an integrative fronto-parietal control network (FPCN). The EES reflects elementary processes that integrate exteroception, proprioception, kinaesthesia, and interoception to establish a physical self-percept, organized at the level of the unconscious (James, 1890; Damasio, 1999; Craig, 2003). The NS describes a self-concept based on reflective and evaluative perception of physical, social, and psychological domains (Christoff et al., 2011). In contrast, EPS comprises higher level percepts acquired through self-specifying, primarily non-judging cognitive processes during present awareness; EPS is thus distinct from the self-related processes of the NS (Gallagher, 2000). While EES, EPS, and NS may be functionally independent, the FPCN generates consistent hub patterns, activating each system differently in practiced tasks and flexibly adapting to novel tasks (Vincent et al., 2008; Cole et al., 2013).

Dimensions of believing and interrelations with mindfulness

Beliefs and disbeliefs are unequivocally mental processes with specific neural correlates, as Sacks and Hirsch (2008): According to the seminal work of Harris et al. (2008), contrasting beliefs and disbeliefs evinced consistently involvement of ventral medial prefrontal cortex (mPFC) using fMRI when subjects assessed written propositions. Interestingly, subjects were quicker to judge statements true than false or undecidable, suggesting that the latter two judgements require more complex information processing. Seitz et al. (2022) proposed three categories of beliefs based on their inherent processual properties: (1) empirical (implicating objects); (2) relational (implicating events); and (3) conceptual (implicating narratives).

These categories reflect varying mental demands and relationships between knowledge and belief. Beliefs are propositional attitudes like desires; at best, they are probabilistic approximations of reality because our sensori-motor and cognitive perceptive systems are limited, as are our predictions of emerging actions (Howlett and Paulus, 2015; Seitz et al., 2016). An example of an empirical belief would be probabilistic modeling of the sensori-motor hand skill of object exploration, which relies on extracting the first three components from a digital data glove. Analyzing the principal component with around 80% variance would enable us to describe finger positions in space over time and thus designate the type of the multifinger task as finger gaiting (Krammer et al., 2020). Structurally, the probabilistic map of the brain lesion (part of a distributed cortical neuronal network) predicted recovery from tactile agnosia vs. persistent disorder over the long-term with 90% accuracy (Abela et al., 2019).

Relational beliefs include percepts of objects and subjects (Seitz et al., 2022). Objects, tools, or interfaces may be integrated because people believe and trust in their usefulness. Eventually, this iterative and embodied process becomes a routine in which use is automatic (Nehaniv et al., 2013). Personal interactions are similarly mediated and stabilized by trust in familiar wordings, manners of speech, and concomitant intimate gestures validated as individuals grow (Seitz et al., 2018). Conceptual beliefs appear in our narratives, often in stories about our past and thoughts about our future, and shape our autobiographical memory (Fivush et al., 2011). Confronted with a conceptual question, subjects decide whether to seek maximal value based on momentary beliefs or explore an issue from several perspectives with the goal of preserving multiple options. The decision to seek maximal value is driven by experience and promises of reward; the action may consolidate our beliefs or make us rigid (Duncan and Peterson, 2014). The decision to preserve options spring from mindfulness, which allows us to better adapt to a concrete situation because it grants us more freedom and can update our beliefs (Langer, 2014).

Believing processes are products of the empirical, relational, and conceptual processes detailed above, and are distinguished by self-relational valuation; they spur action and a learning process that helps us predict errors (Seitz et al., 2018). Multiple factors set the course of a person's believing processes: (1) becoming aware of actions or internal narratives; (2) experiencing agency and ascribing ownership; (3) referring perception to the real world; (4) emotional binding and increasing trust that comes from relying on percepts. When narratives productively use unrealized possibilities, this may raise the risk of counterfactual explanations (Brugger and Graves, 1997). The limitations of the pure third person perspective of classical cognitive science are clear when we examine trusting beliefs. We may trust intentions, behavior, dispositions, and institutions, posing difficulties for an operational definition and modeling trust related judgements (Vidotto et al., 2012).

In contrast, mindfulness is self-generating and self-sustaining, resistant to mindset manipulation (Langer et al., 2010). Mindfulness is distinguished by a mainly non-judgmental behavior, facilitated by a decentred attitude (Shepherd et al., 2016). Decentring makes open-minded acceptance possible and is an essential component of self-awareness. Decentring mediates between mindfulness and positive affect, but not between mindfulness and positive thinking. Mindfulness, however, correlates directly with positive thinking (ben Salem and Karlin, 2022). This suggests that decentring and mindfulness are separable constructs that travel distinct pathways (Gecht et al., 2014). But mindfulness and awareness intertwine and together make it possible for people to perceive thoughts, beliefs, motivations, and feelings clearly (Brown and Ryan, 2003; ben Salem and Karlin, 2022).

Table 1 gives an overview of up-dated personality concepts, which include now experience from a first person perspective and is basic for verification of mindfulness effects on behavior by neuroimaging methods.

TABLE 1 A change of personality concept to integrate the subject's experiential mode of information processing.

(1) Classical dual concept of personality posits two information processing systems in humans: a rationale and an experiential one.
(2) The dilemma of classical cognitive science is: Subjective experience is not accessible to other observers whereas perception of objects and processes are accessible.
(3) The subject exhibits two modes of consciousness: a non-reflective (i.e., a 1st first person perspective) and a reflective (i.e., a 3rd person perspective, ownership of experience).
(4) A probabilistic model posits: the metaphysical self ("I," the subject of experience) vs. the phenomenal self ("me," the object of experience).
(5) A practical approach for a science of consciousness: exact describing the phenomenal-self according to a reflective, methodically guided phenomenological analysis.
(6) Mindfulness and believing interact with living experience and are mutually antagonistic (principle of subjective detachment vs. principle of subjective evaluation).
(7) Here the objectives of the assessment of mental processes are: to differentiate between the manifestations and mechanisms of unconscious EES, conscious EPS and NS.

Searle (1994), Epstein (2003), Gallagher and Varela (2003), Vago and Silbersweig (2012), Chalmers (2013), Choifer (2018), Metzinger (2018), and Wozniak (2018).

Studying mindfulness with neuroimaging using MRI—Concepts and technical aspects

In the early 1990s, functional magnetic resonance imaging (fMRI) made it possible to extensively and non-invasively study cerebral physiology and mental processes, which previously could only be investigated through joint analysis of lesions and disease. The physiological principal underlying fMRI derives from the fact that the magnetic properties of hemoglobin depend on the level of oxygenation in the brain: the BOLD effect (Logothetis and Pfeuffer, 2004). When activated by, e.g., a motor task, oxygen concentration in the related capillary network of stimulated cortical and subcortical areas exceeds normal levels. Brain activation can be compared during active performance and non-performance of a task because fMRI shows blood supply changes in the regions implicated in that task. A time series of individual fMRI scans extend over the course of minutes as the subject cycles through task and control conditions. Most commonly, analyses use a general linear model to make categorical comparisons of the conditions (Friston et al., 1995; Calhoun et al., 2001).

In regions not implicated in a task, we expect brain activity to decrease because sensory modalities are not stimulated (Haxby et al., 1994; Kawashima et al., 1994; Buckner et al., 1996). Deactivation should also be apparent in the frontal and posterior midline cortices (Ghatan et al., 1995; Baker et al., 1996). Andreasen et al. (1995) showed that these areas activated in a memory task and suggested that they were associated with personal reflection. Subsequently, Shulman et al. (1997) and Mazoyer et al. (2001) identified specific brain areas that were more active during passive than during goal-directed task conditions, constituting the "default mode network" (DMN). Gusnard and Raichle (2001) affirmed the functional importance of the passive resting state, proposing that it sustains a stable, unified representation of the individual in their environment: a self-representation. The discovery of the DMN provided significant impulse to explore human cognitive and psychological activity.

The DMN is explored with resting state-fMRI (rs-fMRI), a time series of individual scans taken over the course of minutes. Unlike act-fMRI, rs-fMRI are acquired only in the resting state and are not compared to scans taken in the active state (Biswal, 2012). Commonly, time correlations are computed among the regions of interest captured by brain images to establish functional connectivity. Regions that belong to the DMN are deactivated during act-fMRI studies but show increased activity during periods of reduced interaction with the external world, e.g., rest, sleep, or under anesthesia (Buckner et al., 2008). There is an anticorrelation between the DMN and externally activated networks. The DMN develop in early infancy and is deficient in Alzheimer's disease, autism, and schizophrenia (Buckner et al., 2008). Although we do not yet know its function, the components of DMN were revealed in studies of meditation, self-reflection, perception of prospects, and reflections about others during mentalizing, which is a form of cognitive empathy (Frith and Frith, 2003; Choi-Kain et al., 2008). These studies suggest clear differentiations between reasoning about another person's mental state and affective states shared with another person (empathy associated with distress) or concern for another (compassion), even though these behaviors interact under certain circumstances (Preckel et al., 2018).The two study paradigms, act-fMRI, and rs-fMRI, both acquire time series that extend over periods of minutes. Structural MRI (s-MRI) requires iteratively reconstructing k-space by acquiring signals averaged over minutes. s-MRI uses modern scanners to capture high resolution structural images. Researchers combined sophisticated analysis software with these high-resolution images to develop voxel-based morphometry, allowing them to measure the size of local gray matter in cross-sectional studies and tensor-based morphometry expressed by tensor gradients in longitudinal studies (Ashburner and Friston, 2000; Abela et al., 2014).

The three MRI study paradigms rely on segmenting brain matter into ventricles, white matter, and gray matter (the cortical layer and subcortical nuclei) and spatial standardization of individual brains. The creation of a common stereotactic space makes possible direct comparisons of regional changes in individuals or groups and allows us to study their relation

TABLE 2 Literature search in PubMed [National Library of Medicine, USA].

Assimilation of new settings into the context of mindfulness

Keyword 1	Keyword 2	Keyword 3	Papers [*n*]	Since
Mindfulness			12,008	1985
Mindfulness	Meditation		2,902	2001
Mindfulness	MRI		683	2001
Mindfulness	Meditation	MRI	127	2006
Mindfulness	Awareness	MRI	88	2000
Mindfulness	Believing	MRI	29	2001
Mindfulness	Focused attention	MRI	28	2003
Mindfulness	Mind wandering	MRI	28	2007
Mindfulness	Working memory	MRI	15	2005
Mindfulness	Open monitoring	MRI	4	2001

to behavioral covariates. Long-term changes in local brain volumes over time may indicate brain plasticity, such as might be due to brain lesions or to physical or mental training (Debarnot et al., 2014).

A search of the MEDLINE metadata bank of references and abstracts on life sciences and biomedical topics (National Library of Medicine, USA) *via* PubMed yielded 12,008 publications on mindfulness since 1985 (Table 2) The Table shows the stepwise integration of various mental processes (e.g., meditation, awareness, believing, attention, mind wandering, and working memory) visualized with MRI into the research focus. This expansion of research focus concurred with the detection of the DMN.

Neuroimaging studies of mindfulness—Shaping the brain in parallel with the experience in mindfulness meditation

Using the neuroimaging techniques act-fMRI, rs-fMRI, and s-MRI, described in the previous section, we next present the results of selected original publications that discuss key elements of mindfulness and/or its behavioral covariates observed during naturalistic tasks (Gallagher and Brøsted Sørensen, 2006). Selection was performed with "mindfulness," "focused attention (FA)," "open monitoring (OM)," "mind wandering," "emotional regulation," "magnetic resonance imaging (MRI)" and "default mode network (DMN)." We accommodated our approach to the suggestions of Gallagher and Brøsted Sørensen (2006) to reduce the data and extract the essentials from the observed behavior including its context and, thus, associate the core of the experiential phenomenology with the neuroimaging findings for objectivation. In essence, we categorized first-hand experience of phenomena at a level of abstraction sufficient to allow us to recognize the common properties of phenomenological data and objective data accepted by the sciences. These behavioral categories included task description, task performance, context of the task, explicit or implicit information processing, and experience in mindfulness meditation.

Researchers have characterized subjects' abilities to process emotions during different stages of meditation experience and while exposed to different conditions. Herwig et al. (2010)'s act-fMRI study of meditation-naïve healthy volunteers revealed that BOLD increased in the dorsal mPFC, extending to the superior frontal gyrus, during self-related perception and emotional introspection. At the same time, activity decreased exclusively in the left amygdala during emotional introspection. The unique association of BOLD responses in dorsal mPFC and amygdala during emotional introspection indicated that the phenomenon was independent of voluntary intention. BOLD response within the anterior mPFC during cognitive self-reflection and within the posterior mPFC during emotional introspection correlated inversely with FMI-scale score (Walach et al., 2006), which suggests that subjects with higher mindfulness scores use fewer neural resources. Application of the same study protocol confirmed that emotion-introspection downregulated amygdala activity in depressed patients, supporting its use as mindfulness related treatment (Herwig et al., 2018).

Murakami et al. (2015) continued to explore the relationship between the amygdala and the PFC in a study of unselected healthy subjects presented with images containing emotionally negative content. Subjects used two strategies, voluntary suppression and mindful emotional self-regulation, to cope with these negative images; both strategies reduced negative affect more than natural responses. These strategies suppressed the response of the amygdala, but act-fMRI suggested they each involved different neural systems. During mindful self-regulation, functional connectivity between amygdala and mPFC was prevalent; during voluntary suppression, functional connectivity between amygdala and dorsal lateral prefrontal cortex (lPFC) was prevalent. A post-examination interview indicated that mindful introspection was accompanied by more reliable self-monitoring. Kral et al. (2018) showed that amygdala response to negative emotional stimuli decreased in more experienced meditators, while short-term training in a mindfulness-based stress reduction course did not have the same effect. In an act-fMRI study by Lebois et al. (2015), healthy subjects who had not previously meditated were taught two strategies for disengaging from one sentence scenarios (i.e., stressful vs. non-stressful) projected on a screen. The first strategy was mindful attention (a decentring attitude) and the second was immersion in the scenario. Those who paid mindful attention showed less neural activity in the subgenual ACC, ventral ACC, ventral mPFC and medial orbito-frontal cortex during exposure to stressful scenarios that those who immersed

themselves in the scenario. Three day intensive mindfulness meditation training intervention has been effective in reversing resting state functional connectivity between amygdala and subgenual ACC which associates previously perceived stress (Taren et al., 2014).

Reappraisal and acceptance can also be used to cope with emotionally negative experiences. Unlike acceptance, reappraisal is an elaborate cognitive practice in which one iteratively reinterprets negative experiences so that they eventually cease eliciting negative affect (Gross, 2001). In healthy subjects confronted with sad images, both strategies effectively regulated negative emotions better than no strategy; reappraisal was more effective than acceptance (Smoski et al., 2015). This act-fMRI study found the right frontal pole and medial frontal cortex were activated during acceptance. The left insula and precentral gyrus were activated during reappraisal. Opialla et al. (2015) observed activation in the left ventral and dorsal lPFC, supramarginal gyrus, and insula increased with mindfulness more than with cognitive re-appraisal during cued expectation of negative stimuli, but not during perception. Thus, initiation of a mindful state may engage more neural resources, specifically in the expectation phase of meditation-naïve subjects. Subjects who had major depressive disorders had less activity in their ventral medial PFC during mindful acceptance, a predictive sign of depressive relapse, and less activity in cognitive control regions like the paracingulate area, which may influence the ability to adapt emotional responses (Shackman et al., 2011).

Modinos et al. (2010) used an act-fMRI study design to investigate healthy subjects and determine the relationship between mean activation in dorsal mPFC during reappraisal and mindfulness traits as determined by KIMS. Subjects were taught a reappraisal strategy and then confronted with neutral and negative images during a fMRI task. Their dorsal mPFC was activated when they successfully reappraised the negative image while their amygdala deactivated. The positive association of mean activation in dorsal mPFC with mindfulness traits according to KIMS was dependent mainly on the subscale "act with awareness." Other studies unrelated to mindfulness found that the degree of ventral mPFC activation was positively related to self-referential evaluations (Gusnard and Raichle, 2001; Northoff and Bermpohl, 2004) and evaluations of emotional stimuli affecting others (Frith and Frith, 2003).

Some effects of meditation practice were touched on above, and these and other effects have been the subject of extensive investigations. In an act-fMRI study, Farb et al. (2007) used a word perception task to show that experienced meditators were more likely to enter a state of experiential self-awareness than naïve subjects. Experienced subjects were trained daily in an 8-week mindfulness-based stress reduction course. They learned to discriminate between experiential and narrative forms of self-awareness, while untrained naïve subjects could not differentiate the two and tended to the narrative mode. All subjects were asked to engage in self-awareness while they were briefly shown a validated, randomized list of words chosen to elicit either positive or negative emotions (Anderson, 1968). BOLD response in both the ventral and dorsal mPFC markedly decreased in experienced subjects in the experiential mode. At the same time, activation shifted from midline structures to right prefrontal-lateral areas and the anterior insular cortex (AIC). The authors proposed this shift might be caused by consolidation of a decentred attitude and diminished self-referential neural processing. Farb et al. (2013) extended this study, employing the same subjects but changing the test paradigm. The revised study design included three conditions: interoceptive attention; word perception while refraining from cognitive or emotional response; and recognition of word repetitions. The state of interoceptive attention in experienced meditators deactivated the dorsal mPFC and activated the right posterior insula. Concurrent activation of the posterior and anterior insular cortex also indicated that the internal insular structure was reorganized. This activity pattern, including dorsal mPFC, might indicate the substratum in which tonic activity is preserved in the AIC during externally focused attention. The greater functional connectivity between posterior and anterior insula may enable subjects to integrate simultaneously interoceptive and exteroceptive processing. The crucial active involvement of the AIC during interoception and its importance for interoceptive accuracy, e.g., toward sensing the breathing rhythm, were established in recent fMRI experiments by Wang et al. (2019). Recently (Lenhart et al., 2020) reported findings similar to those of Farb et al. (2013) in a longitudinal study of gray matter changes in the course of a 7 weeks FA meditation training. They found increases in the AIC, the caudate nucleus and the frontal cortices, decreases in the parieto-temporal areas and the parahippocampal gyrus (PHG) as well as fractional anisotropy alterations adjacent to right hippocampus (HIC) and basal ganglia. Most important are the contributions of Santarnecchi et al. (2021) using rs-fMRI and determining mindfulness induced functional connectivity in the right putamen, cerebellum and anterior insula after an 8 week MBSR training. The prominent findings were the effective connectivity patterns between ACC, putamen on both sides and right cerebellum and the differential response of executive and somatosensory putaminal subregions within this network, exerting a modulatory functional impact both on orbitofrontal cortex and cerebellum.

A four-condition act-fMRI by Brewer et al. (2011) revealed that the activation pattern of experienced mediators was independent of three test conditions: attention to respiration; loving kindness; and neutral awareness. In contrast to controls deactivation of the posterior cingulate cortex (PCC) and mPFC, two constituents of the DMN, were observed in experienced meditators as part of this network during the three test conditions relative to baseline. Experienced mediators also exhibited more connectivity in PCC and areas involved in conflict monitoring, cognitive control and working memory

(the dorsal lPFC and ACC) under all test conditions (Mansouri et al., 2009). Hasenkamp et al. (2012) and (Hasenkamp and Barsalou, 2012) used rs-fMRI to monitor cognitive processes in experienced meditators and a control group of subjects with little meditative experience. The authors differentiated four consecutive phases of a naturalistic cycle during self-monitoring: mind wandering, awareness of mind wandering, attention shifting, and re-established sustained attention. These form a theoretical model of cognitive fluctuations during a FA session. Experienced meditators were distinguished by diminished activity in a cluster that involved ventral mPFC and ACC when they shifted and restored sustained attention. Dorsal lPFC plays a key role in preserving FA (Hasenkamp et al., 2012). Significant differences between experienced and inexperienced mediators have also emerged in an analysis of functional connectivity. This analysis delineated a salient network involving DMN at ventral mPFC and bilateral PCC during mind wandering and involving executive network at dosal lPFC while sustained attention to breathing was restored. When experienced mediators paid sustained attention, connections between the right anterior insula, left dorsal IPFC, mid cingulate gyrus and right dorsal IPFC increased as did connections between a PFC/ACC cluster and bilateral inferior parietal lobules, while connectivity between a ventral mPFC/ACC cluster and the left PCC decreased (Hasenkamp and Barsalou, 2012). Similarly, in a rs-fMRI study of subjects who participated in a 4-day intensive meditation course resulting in sustained resilience for 3 months in contrast to controls in a relaxation retreat, Kwak et al. (2019) observed an increase in resting state functional connectivity between the dorsal mPFC and rostral ACC. A rs-fMRI study of healthy elderly subjects revealed that MAAS-scores and two constituents of the DMN (PCC and precuneus) correlated positively (Shaurya Prakash et al., 2013). The dorsal area of the PCC delineated most prominently in the correlation may interface between the resting state network and the network regulating cognitive control (Leech et al., 2011). The precuneus fulfills multiple functions, among them maintaining open monitoring (Gusnard et al., 2001).

The mind wandering phase of the monitoring cycle complements mindfulness. In an act-fMRI study of healthy volunteers, Mason et al. (2007a) explored the occurrence of stimulus independent thoughts (SIT) during verbal and visuo-spatial tests of working memory. They found a voxel-wise correlation between the tendency to daydream (assessed on the Imaginal Process Inventory-scale) (Singer and Antrobus, 1970) and constituents of the resting state network, e.g., the ventral and dorsal mPFC, posterior cingulate cortex and precuneus. They suggested as SIT accumulates during the tasks, the transition to mind wandering is inevitable. Mind wandering hinders mindful perception and eventually elicits ruminations about past and future. In an fMRI study of neural recruitment in both the default mode and executive networks, Christoff et al. (2009) suggested that mind wandering is most pronounced when meta-awareness is absent. Based on rs-fMRI, Wang et al. (2014) defined eleven nodes of the DMN based on their positive functional connectivity to PCC: PCC (1); mPFC (2); superior frontal gryrus on both sides (3, 4); lateral parietal cortex on both sides, LPC (5, 6) (i.e., BA 39); lateral temporal cortex on both sides, LTC (7, 8); PHG on both sides (9, 10); and thalamus, TH (11). Based on MAAS-scores, the link between thalamus and PCC most closely correlated with mindfulness. Nodal properties of the thalamus exhibited weak but significant negative correlations with these scores.

Mindfulness both activates distinct regions of the brain and induces morphological plasticity in the long-term. Using high resolution s-MRI, Murakami et al. (2012) established a correlation between the FFMQ-scale and the volume of the right insula and amygdala in healthy subjects. They suggest that volume increases in these structures might comprise a module in which the right insula facilitates physical interoception and the amygdala facilitates emotional response. Chronic emotional stress might also increase the volume of the amygdala, as Gianaros et al. (2008) observed. In a second s-MRI study of 247 college students with no previous experience of meditation, Lu et al. (2014) found MAAS scores correlated with the volume of gray matter in areas of the DMN and attention networks. The PCC on both sides, the left orbito-frontal cortices (OFC) and the right HIC/amygdala correlated negatively, while the dorsal ACC on both sides correlated positively. The positive correlation between MAAS score and dorsal ACC volume indicates that the dorsal ACC plays a role in sustaining attention and thus in conscious awareness. The negative correlation between MAAS score and PCC volume is consistent with the decrease in self-related thinking in more mindful students and the negative correlation between MAAS score and OFC and HIC/amygdala volume is consistent with reduced emotional responsiveness. In a third s-MRI publication, Zhuang et al. (2017) explored the disposition to mindfulness in a large group of young adults with no experience meditating. The authors connected MAAS and FFMQ scores to brain volumes and surface areas and found MAAS scores and gray matter volumes significantly correlated with the volume of the right precuneus, the surface area of the right dorsal lPFC (Brodmann area, BA, 46), the right inferior parietal lobule, IPL (BA40), and the left superior prefrontal cortex (BA 9). They also significantly correlated with the FFMQ items that comprised the category "describing." These findings are consistent with increased self-awareness in more mindful young adults.

Several studies examined the effects of long-term meditation experience on brain morphology and function. Lazar et al. (2005) found focused attention significantly increased thickness in the right anterior insula and prefrontal cortex (BA 9 and 10); increases were less significant in the somatosensory, auditory, and visual cortices. Luders et al. (2012)'s cross-sectional study

compared experienced meditators (mean ± SD 19.8 ± 11.4 years of experience) to healthy controls and found that in the experienced group mean curvature increased, suggesting increased cortical gyrification, maximal in the right anterior dorsal insula. Their prediction was based on the hypothesis that group differences and/or correlations would be most pronounced in cortical regions known to increase in volume in meditators, e.g., in the right anterior insula (Lazar et al., 2005; Hölzel et al., 2008). They also found that curvatures increased to a lesser extent in the left anterior dorsal insula, left precentral gyrus, right fusiform gyrus, and right cuneus. The authors emphasized the key role the anterior insula plays in long-term meditation. In a complementary rs-fMRI study, Taylor et al. (2013) found that the functional correlation between right IPL (BA 39) and dorsal mPFC (BA 10) was stronger in experienced mediators than in novices. This pattern of interconnected nodes suggests that long-term meditation may improve global attention rather than mindfulness, specifically since the parietal cortex is involved in working memory and affects visuo-spatial performance (Courtney et al., 1996). Recently, Fujino et al. (2018) discovered specific subcortical—cortical interactions in experienced meditators. Functional connectivity from the striatum to the posterior cingulate cortex diminished during FA meditation and during OM meditation, but functional connectivity from the ventral striatum to the retrosplenial cortex, which maintains memory function in the DMN, diminished only during OM-meditation. For the first time, this segregation from memory function was substantiated with neuroimaging, revealing a mechanism of detachment from self-relational thoughts.

When they compared different forms of meditation to visual stimulation in long-term meditators using fMRI, Josipovic et al. (2011) found anticorrelation between the task-positive extrinsic (the visual system) and the task-negative intrinsic (the DMN system) decreased during non-dual awareness (NDA) in the Tibetan tradition when referencing to anticorrelation observed in FA meditation. Most important, they found no differences between conditions in the modulation of brain activity in either network. NDA likely differs conceptually from FA and OM meditation, since NDA is context-oriented and FA and OM meditation are content-oriented (Josipovic, 2014).

Table 3 provides a summary of behavioral domains appropriate to the time dependent expertise in mindfulness experience, the observed neurophenomenology in the according tasks and the putative areas associated with the observed neurophenomenology. Figure 1 shows the center of gravity of these areas involved as they relate to DMN (dorsal and ventral mPFC, PCC, IPL) and the insular cortex. The underlying individual studies and the associated areas involved within DMN and insular cortex are summarized in Supplementary Tables S3, S4.

Discussion: Deducing the neurobiological underpinnings of mindfulness from brain imaging—A conceptual approach

Specific brain areas reflect mindfulness cultivated in FA- and OM- meditation training. These areas reflect the degree of experience meditators acquire as they practice and they specifically involve the EES, EPS, NS and FPCN systems, cf (Vago and Silbersweig, 2012).

Self-relational processes and the midline structures of hemispheres

Gallagher (2000) describes the implications for cognitive science, positing in his philosophical conceptions of the self that people have a minimal embodied self-representing consciousness as an immediate subject of experience, existing in the present. He also posits that people possess a supplementary narrative self—a self-image that includes a past and future, inherent in the stories they formulate about themselves. Kyselo (2014) commented on these two aspects of the self, emphasizing that social existence is organized in terms of back-and-forth between social distinction and participation processes. In their view, the body becomes the mediator of these processes. The subject's narrative relies on self-relational processing that includes implicit subjective feelings and explicit cognitive thoughts, each of which are mediated by a task-negative network of cortical midline structures (Northoff et al., 2006) that comprise the ventral-medial prefrontal cortex (ventral mPFC), pre- and subgenual ACC (preACC, sgACC), posterior cingulate cortex (PCC), retrosplenial cortex (RSP), PHG and HIC.

The activity in the ventral mPFC decreases when subjects psychologically distance themselves from self-representations. Inversely activity in the ventral mPFC increases when personal values and self-related thoughts are involved (D'Argembeau, 2013). The ventral mPFC allows subjects to incorporate the interests of the self into an episodic event of the past while the HIC ensures they subject can recall this even in detail (Kurczek et al., 2015). This discovery may help us understand future thinking: patients with ventral mPFC lesions could not describe events in their own future in any more detail than they could describe events that happened to other people in the past (Verfaellie et al., 2019).

Garrison et al. (2013) took a neurophenomenological approach to studying undistracted awareness and effortless doing, and found they were associated with PCC deactivation, while distracted awareness and controlling were associated with PCC activation. Coincident subjective experience during these two antagonistic mental conditions (meditative vs. self-related)

TABLE 3 The self-specifying process in mindfulness from subjects naïve in mindfulness to expert status proficient in the metacognitive skill.

Observed neurophenomenology is dependent on duration of mindfulness training				
Status	**Duration of meditation training**	**Behavioral domain**	**Neurophenomenology of tasks**	**Putative structures related to neurophenomenology**
No experience and novices	0	Mind wandering	Activity independent thoughts	FC bw mPFC, PCC and precuneus[a]
	On a waitlist for meditation courses	Emotion control	Voluntary suppression	Amygdala connected to dlPFC[b]
		Emotion control	Mindful self-regulation	Amygdala connected to dmPFC[b]
↓		Emotion control	Introspection vs. self-reflection	SLF (dmPFC) ↑ and amygdala ↓[c]
		Mindful disposition	Expression of mindfulness traits according validated scales	FC bw PCC and Precuneus[d]
		Mindful disposition		FC bw thalamus and PCC (structure)[e]
		Focused attention	"Distracted" awareness*	BOLD in DMN (↓), mostly in dmPFC ↓[f]
Initial experience	Months	Open monitoring	Experiential vs. narrative focus	dorsal mPFC ↓, ventral mPFC ↓[g]
↓		Open monitoring	Integration of IC and EC	dmPFC interacts across different conditions[g]
Advanced experience	Years	Open monitoring	Self-relational detachment	FC bw striatum and retrosplenial cortex[h]
		Steering mindful state	Recognition of lapse into MW	dorsal ACC, bilateral anterior insula ↑[i]
↓		Steering mindful state	Shifting from MW to FA	mPFC/ACC in high meditation practice ↓↓[j]
		Global attention	Moment-to-moment awareness	FC of dorsal mPFC L to IPL R ↑[k]
Experts	> 5 years	Focused attention	"Undistracted" awareness*	PCC ↓↓[l]
		Metacognition	Control across meditation states	Co-activation of mPFC, PCC, dorsal ACC, dlPFC[m]
		Non-dual awareness	Access to context-oriented info	Synergism between DMN and sensory networks, e.g., a visual network processing[n]

IC, interoception; EC, exteroception; FC, functional connectivity; MW, mind wandering; FA, focused attention; bw, between.
Appropriate references: [a]Mason et al. (2007a); [b]Murakami et al. (2015); [c]Herwig et al. (2010); [d]Shaurya Prakash et al. (2013); [e]Wang et al. (2014); [f]Scheibner et al. (2017); [g]Farb et al. (2013); [h]Fujino et al. (2018); [i]Hasenkamp et al. (2012); [j]Hasenkamp and Barsalou (2012); [k]Taylor et al. (2013); [l]Garrison et al. (2013); [m]Brewer et al. (2011); [n]Josipovic et al. (2011).
*"Distracted" awareness means switching between FA, MW and refocusing whereas "undistracted" awareness means longer phases of undisturbed focusing.

were quite different and seemingly specific (Garrison et al., 2013). On the basis of its connections, it seems the RSP is uniquely positioned to translate between the world-centered domain, including perirhinal gyrus, HIC and PHG, and the self-centered world of the medial parietal cortex (Vann et al., 2009). These are the areas involved in the phases of mind wandering (Mason et al., 2007a; Hasenkamp and Barsalou, 2012).

Analogous mechanisms have been uncovered in the course of believing processes, displaying activation patterns while subjects evaluated self-related interests or preferences. Independent of testable and non-testable beliefs, main effects of certainty were evident in the involvement of a midline neuronal network encompassing the left mPFC at intermediate z-level, caudate and PCG, and right superior temporal lobe in the neighborhood of temporo-parietal junction (TPJ). Certainty of a non-testable proposition, a strong belief, activated the left insula (Howlett and Paulus, 2015). Common areas engaged in false belief reasoning and visual perspective taking, which is a precondition for assessing the perspective of another subject while mentalizing, are evident in the left angular gyrus; these areas include the temporo-parietal junction, and the left medial occipital gyrus (Schurz et al., 2013). Incongruent mental states (false beliefs and unfulfilled desires) and congruent mental states also significantly increase the involvement of PCC/RSC in processing unfulfilled desires, while the same level of involvement is not shown for true beliefs (Abraham et al., 2010).

Unconscious self-specifying processes

EES integrates implicit activity of a subject with prevalent automatic responses to extero- and interoception (Aspell et al., 2013), functional also outside of the focus of awareness (Roeser and Peck, 2009). Enactivism is primarily an implicit ongoing iterative process that helps the subject create a world of meaning through interaction with the environment, including other subjects. Enactivism is independent of logic presumptions and does not rely on representations (Nehaniv et al., 2013; de Haan, 2021). The process of enactivism is supported by embodiment, which structurally couples the subject with the world and results in non-conscious embodied actions (Nehaniv et al., 2013;

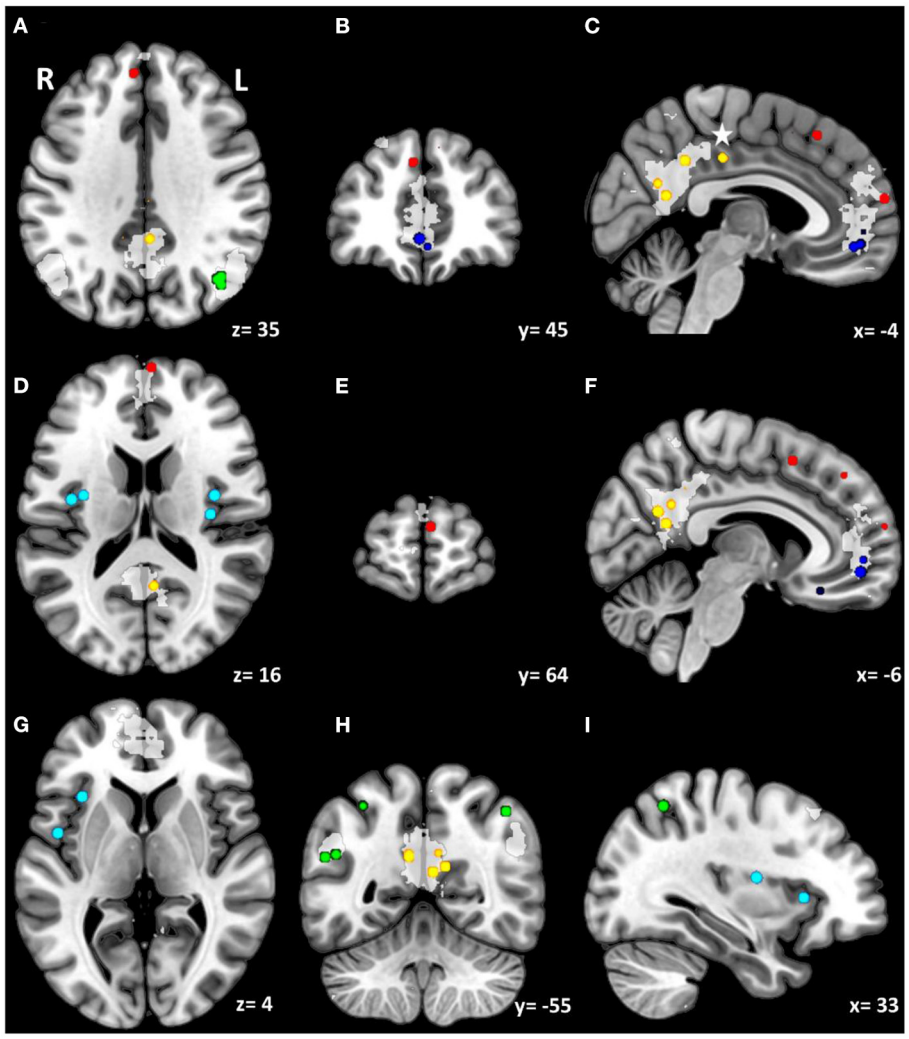


FIGURE 1
Involvement of default mode network constituents as well as the insular cortex. The Figure indicates centers of gravity of cortical involvement observed in selected neuroimaging studies detailed in Supplementary Table S3, integrated into the standard MNI 152 template. For according MNI coordinates see Supplementary Table S4. Enclosed is furthermore an automated topic-based meta-analysis using the term "DMN" in article abstracts provided by https://neurosynth.org/analyses/terms/dmn/ for comparison purposes. The light-gray areas superposed on the anatomical slices delineate zones preferentially associated with the term in 366 neuroimaging studies with an expected FDR < 0.01. The red and dark blue spheres indicate the involvement of dorsal and ventral mPFC, and the yellow spheres indicate PCC involvement. **(A–F)** Self-specifying processes involve the dorsal mPFC whereas self-relational processes involve ventral mPFC, together with pre- and subgenual ACC, as well as PCC and retrosplenial cortex. Please note: Involvement of mPFC at the level of superior frontal gyrus is predominant in subjects of no meditation experience suggesting voluntary effort during a task. **(C,F)** The dorsal area of the PCC, marked by a star, may be a separate compartment: an interface between the resting state network and the network regulating cognitive control (Leech et al., 2011). Ad insular cortex (light blue): Proximal insular cortex is a primary interoceptive center with distinct homeostatic functions **(D,G,I)**, whereas dorsal anterior insular cortex has been shown to support explicit interoceptive attention **(I)** (Wang et al., 2019). **(A,H,I)** IPL (green spheres) at its posterior part (angular gyrus) is related to the DMN, whereas at its anterior part (supramarginal gyrus) to the FPCN.

Izmirli, 2014; Varela et al., 2016). As detailed above, EES is linked with the NS by the midline brain structures but distinguished from the NS by the underlying task-positive network.

In contrast, active enactive experience involves the subcortical-level midbrain nuclei, superior colliculi, medio-dorsal and ventral-posterior thalamus, pulvinar and dorsal striatum, and the cortical level proximal insula, premotor, and sensory association areas (Damasio, 1999; Craig, 2003). Activation of the proximal insular cortex is prototypical for afferents that transmit physiological information about distinct homeostatic sensory modalities. Proximal insular cortex activation is associated with an equivalent homeostatic emotion that engenders distinct body feelings and preserves physiological balance (Craig, 2004). In humans, an increasing proximal-to

mid-to anterior pattern parallels integration of distinct sensory information and contextual affective contents [Stephan et al., 2003; Bud Craig, 2009]. This homeostatic processing provides the subject with diverse information (homeostatic motor functions, and environmental, hedonic, motivational social and cognitive conditions) that is integrated into a meta-representation within AIC, and with simultaneous co-activation in the ACC. This information, merging from various sources into a meta-representation, creates an emotional moment characterized by a specific feeling and an associated emotion (Craig, 2009). Sterzer and Kleinschmidt (2010) discuss the role of AIC in perceptual processes, asking if the AIC supports awareness of the immediate moment in a state of a subject's heightened alertness. Farb et al. (2007) validated the active involvement of a right lateralized network including lPFC, AIC, secondary sensory cortex, and inferior parietal lobule, which suggests experiential focus centers on the present in trained meditators. Mindfulness meditators may perceive a slowing of time in the present based on their ability to focus more strongly on sensory experiences and to be more strongly aware of feelings and of body states (Wittmann and Schmidt, 2014). The argument that AIC is integrated into perception of time intervals in the range of seconds to sub-seconds is supported by fMRI task results (Livesey et al., 2007). In the context of this section the involvement of large-scale networks should be noted as reported in recent papers, including subcortical gray and white matter, brain stem and cerebellum (Tang et al., 2015; Lenhart et al., 2020; Santarnecchi et al., 2021).

Conscious self-specifying processes

Arising from unconscious information processing of EES, subjects develop a self-as-subject or a minimal self that is not taken as an intentional object; instead, it acquires knowledge from a first person perspective (Gallagher, 2000; Legrand, 2007). At the level of the experiential phenomenological self, individuals consciously perceive information of subjective content, however these percepts cannot be transformed into third person data through traditionally valid scientific procedures (Gallagher and Brøsted Sørensen, 2006). To prevent methodological biases (Gallagher, 1997; Gallagher and Varela, 2003) suggested a framework of reflective, methodically guided phenomenological analysis of behavior to get information about the phenomenal self—the "me." The task-positive mental processes of experiential self are associated with attention and anticorrelated to the task-negative mental processes of NS, which are associated with long-term memory (Buckner and Carroll, 2007).

The synopsis of conscious self-specifying vs. self-related processes together with functional and structural neuroimaging studies yields main findings as detailed in the section "Neuroimaging studies of mindfulness—shaping the brain in parallel with the experience in mindfulness meditation." Functional connectivity between PCC and thalamus plays a dominant role in self-specifying processes. Connectivity is inversely proportional to mindfulness (Wang et al., 2014). Dorsal mPFC is a main hub of mindful disposition and behavior over a wide range of experience in meditation training (Farb et al., 2007, 2013; Herwig et al., 2010; Modinos et al., 2010; Kral et al., 2018). Specific strategies against negative emotions are clear. The right dorsal mPFC correlates with the left amygdala (Murakami et al., 2012) when viewers see pictures with negative valence; dispositional mindfulness correlates with mindful disposition, based on KIMS. When reappraisal for anticipating negative emotions was directly compared to voluntary suppression, the pathway for anticipating negative emotions was through the right dorsal mPFC to left amygdala, and the pathway for voluntary suppression was through the right dorsal lPFC and left precuneus to the left amygdala (Murakami et al., 2015).

The activation pattern in emerging daydreaming changes from the dorsal mPFC to the ventral mPFC and PCC when the thinker transitions to self-relational thoughts (Mason et al., 2007a). In elderly individuals, connectivity between posterior PCC and medial precuneus cortex correlates with mindful traits, which may reflect the multiple functions of PCC at this site, some of which may specifically maintain open monitoring (Gusnard et al., 2001; Shaurya Prakash et al., 2013). When ruminative thoughts simulated by immersion were compared to disengaging by subjective decentring, researchers found distinct spatial patterns in the structures involved for each condition in non-meditative individuals. Mental immersion involved brain areas that reflected bodily and experiential self-relation, e.g., ventral mPFC, mOFC, vACC, sgAC. Mindful intention involved areas that indicated perspective shifting, e.g., dorsal mPFC, IPL, including angular gyrus (Lebois et al., 2015).

In an experiential vs. narrative test paradigm, an 8-week mindful meditation training course reduced BOLD in the dorsal mPFC to levels lower than those found in meditation novices in Farb et al. (2007). Developing interoceptive attention and mindfulness training evoked greater activity in the anterior insula in experienced meditators (Craig, 2002; Farb et al., 2013). Hasenkamp et al. (2012) found the involvement of dorsal ACC and bilateral AI enhanced when the subject became aware of lapse into mind wandering. Hasenkamp and Barsalou (2012) found the involvement of ventral mPFC/orbitofrontal cortex diminished when the subject shifted from mind wandering to focused attention. In novices this switching between mindful attention, mind wandering and refocusing causes distracted awareness associated with diminished activities within constituents of DMN (Scheibner et al., 2017). Long-term meditators exhibit fundamentally functional changes in DMN connections (Taylor et al., 2013). Very experienced meditators achieved in a FA-task the level of one-point concentration providing them undistracted awareness

associated with pronounced activity decrease in PCC (Garrison et al., 2013). Strong connections were evident between dorsal mPFC and R IPL (most likely corresponding to angular gyrus), precuneus/PCC and R IPL, and R IPL and L IPL (Taylor et al., 2013), which suggest enhanced functional working memory and attention, and diminished self-relational processing (Culham and Kanwisher, 2001; Northoff and Bermpohl, 2004; van Buuren et al., 2010). In masters of introspection, awareness and emotional control, the dorsal-anterior insula was the site of an increase in global maximum gyrification, suggesting this area plays a key role in integrating autonomic, affective, and cognitive processes (Luders et al., 2012). We can distinguish complex NDA meditation from FA and probably also OM meditation because extrinsic networks processing experiences related to the environment and intrinsic networks processing experiences related to interoception are increasingly synergistic in meditators proficient in NDA meditation than competitive (Josipovic, 2014).

Learning processes

Meditation is a form of mental training to acquire the basic prerequisites for maintaining a mindful disposition. In a meta-analysis of 78 functional neuroimaging (fMRI and PET) studies (Fox et al., 2016) found specific but diverging patterns of activations and deactivations when comparing FA, mantra recitation, OM and compassion/loving kindness meditation. Peak activation likelihood estimate (ALE) was given in FA and OM meditation, we were focusing on according to selection criteria for the review: In FA peak values for activations involved left premotor cortex, supplementary motor area, right putamen/lateral globus pallidus, right fusiform gyrus, right cuneus and left precuneus, and peak value for deactivation left anterior insula; in OM peak values for activations involved right anterior insula, right parieto-occipital sulcus and right somatosensory cortices/inferior parietal lobule. In a recent brain theory of meditation (Raffone et al., 2019) suggest a left-brain dominance for top-down regulation in FA meditation and a predominant cognitive and emotional processing in right anterior areas such as the anterior insula connected with the homotopic left hemispheres *via* the frontal parts of corpus callosum. The authors differentiate the mechanism for optimized use of brain resources in FA and OM, through reduction of firing neurons in the former and through tuning the communication between widespread neurons with higher firing rates in a given time window in the latter. Hernández et al. (2018) delineated an ultimate goal of long-term meditators in Sahaja Yoga Meditation tradition: The capacity to maintain a state of mental silence was based on a larger gray matter volume in right anterior cingulate cortex/medial PFC, while performance during scanning evoked increased functional connectivity of this region with bilateral AIC, and decreased functional connectivity with right thalamus/PHG.

In the transition phase between unconscious and conscious processes, contemplative practices may foster attention, emotion regulation, and introspection. These practices may eventually cultivate the habitual patterns of thoughts and beliefs NS provides and establish a mindful disposition governed by the EES (Vago, 2014; Seitz et al., 2016). Technology-mediated mindful intervention studies that used electroencephalographic frequency data to provide the user with real-time acoustic feedback, provide preliminary evidence that mindfulness effectively promotes conscious access to implicit information (Balconi et al., 2017). The evidence from electrophysiological observations is striking. For example, in subjects practicing mindfulness, we see the amplitude of the late positive potential (LPP) decreased after only 400 ms after viewing negative images. Temporo-parietal positivity associated with identification, evaluation, and labeling of the visual stimulus occurs between 600 and 1,000 ms (Brown et al., 2013). Consistent with these findings, mindful intervention during initial observation of negative pictures induced an alternative pathway in which the dorsal mPFC was involved. Late voluntary suppression of the full-blown affect mainly involved the dorsal lPFC (Murakami et al., 2015). The predominant pattern in mindful intervention might illustrate a transition from conceptual to non-conceptual awareness, reducing habitual evaluative processing and involving other areas like the thalamus, insula, sensory, and motor regions (Craig, 2003; Farb et al., 2013). Increased conscious awareness at the somatic and mental level may couple the sensory system to the organism and the environment and at the social level provide more participation (Varela et al., 2016).

Self-specifying and self-relational processes involve the cortical midline structures of DMN in distinct and partly antagonistically manners. These divergences reflect different behavioral levels of concepts of the self (e.g., sensory processing, self-referential processing, higher order processing), which interact in both bottom-up and top-down directions (Northoff et al., 2006).These process dynamics and their mutable participation in mentation shape DMN compartments to cognitive and contextual domains and influence their interaction. This influence is reflected in the evolution of functional and structural cortical patterns in the continuum from subjects naïve in meditation to subjects with long-term meditation experience (Josipovic et al., 2011; Josipovic, 2014). Patterns may change over the lifespan. When (Crane et al., 2020) explored links between personality traits in older people and cognitive performance and the default mode network, they found open perception was associated with three nodes: mPFC; middle frontal gyrus; and dorsal PCC, which may correspond to area 7 m outside of the DMN proper (Vogt et al., 2006; Leech et al., 2011; Shaurya Prakash et al., 2013).

Conclusions

Mindfulness is set by the immediate subject of experience, unextended in time. This is different from the narrative of individuals—a self-image with a past and a future. The immediacy of mindfulness initiates self-specifying processes, primarily at the unconscious level of enactivism and embodiment, and secondarily at the conscious level of experiential phenomenological awareness. Necessary pre-conditions are competence to pay sustained attention, detachment from self-relational thoughts and preferences, and a non-aversive attitude. Practiced over the long-term, mindfulness will improve individual and social wellbeing. As meta-cognitive skill it enables the subject to monitor perception and behavior.

Critical questions remain to be answered. Subjects of varying meditative experience exhibited considerably variable cortical sites of co-activations, so we must clarify the role the dorsal mPFC plays in mindful tasks within the extending area of the midline prefrontal cortex. We might be able to elucidate the exact structural and functional segregation of ventral mPFC by connecting activation likelihood estimates (ALE) of neuroimaging meta-data to specific behavioral paradigm classes of assigned tasks (Bzdok et al., 2013). We also need to understand how the diverse mindful traits are assigned to common or diverse neural networks, and to learn more about how mindfulness increases the capacity of a meditator's working memory (Jha et al., 2010; van Vugt and Jha, 2011; Mrazek et al., 2013). How is effortless attention differentiated from forceful cortical control mechanisms, especially when we perform demanding naturalistic tasks (Gallagher and Brøsted Sørensen, 2006; Jensen et al., 2012; Raffone et al., 2019)? We need to know why segregation of resting state networks seems to decrease processing speed in older subjects when constituents of fronto-parietal control network are affected (Malagurski et al., 2020). Finally, the modified resting-state in long-term meditators may affect mind wandering since the mindfulness may evolve in ways that alter in-parallel self-relational thoughts and induce a more positive mood (Vago and Zeidan, 2016).

Implementing the neuroimaging techniques s-MRI, act-fMRI, and especially rs-fMRI, was a major step forward. These techniques help us understand the dynamic processes underlying mindfulness, follow the process of learning the meta-cognitive skill of mindfulness from early to long-term experience in meditation, and delineate the governing structure of the DMN. The main constituents of DMN are the dorsal mPFC, ventral mPFC, and PCG, which differentially interact depending on the subject's experience in meditation. The midline structures of dorsal mPFC, ventral mPFC, and PCG are antagonistic to self-specifying and self-relational processes, so they allow approximate discrimination in-between. AIC is a meta-representation for sensory perception that integrates both interoception (the self-centered world) and external perception (the world-centered domain). Brain volume changes may indicate brain plasticity, mediated by mental training over the long-term.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

BW conceived and wrote the systematic review, conducted the search, and screened the titles, abstracts, and full texts of the paper.

Funding

This paper was funded by Dr. Rüdiger Seitz, via the Volkswagen Foundation, Siemens Healthineers, and the Betz Foundation. Siemens Healthineers was not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

Acknowledgments

I would like to thank Dr. Kali Tal, University of Bern, for proof reading of manuscript and reviewing grammar and style.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnbeh.2022.928522/full#supplementary-material

References

Abdoun, O., Zorn, J., Poletti, S., Fucci, E., and Lutz, A. (2019). Training novice practitioners to reliably report their meditation experience using shared phenomenological dimensions. *Conscious. Cogn.* 68, 57–72. doi: 10.1016/j.concog.2019.01.004

Abela, E., Missimer, J. H., Pastore-Wapp, M., Krammer, W., Wiest, R., and Weder, B. J. (2019). Early prediction of long-term tactile object recognition performance after sensorimotor stroke. *Cortex* 115, 264–279. doi: 10.1016/j.cortex.2019.01.018

Abela, E., Seiler, A., Missimer, J. H., Federspiel, A., Hess, C. W., Sturzenegger, M., et al. (2014). Grey matter volumetric changes related to recovery from hand paresis after cortical sensorimotor stroke. *Brain Struct. Funct.* 220, 2533–2550. doi: 10.1007/s00429-014-0804-y

Abraham, A., Rakoczy, H., Werning, M., von Cramon, D. Y., and Schubotz, R. I. (2010). Matching mind to world and vice versa: Functional dissociations between belief and desire mental state processing. *Soc. Neurosci.* 5, 1–18. doi: 10.1080/17470910903166853

Ainsworth, B., Eddershaw, R., Meron, D., Baldwin, D. S., and Garner, M. (2013). The effect of focused attention and open monitoring meditation on attention network function in healthy volunteers. *Psychiatry Res.* 210, 1226–1231. doi: 10.1016/j.psychres.2013.09.002

Anderson, N. H. (1968). Likableness ratings of 555 personality-trait words. *J. Personal. Soc. Psychol.* 9, 272–279.

Andreasen, N. C., O'Leary, D. S., Arndt, S., Cizadlo, T., Hurtig, R., Rezai, K., et al. (1995). Short-term and long-term verbal memory: a positron emission tomography study. *Proc. Nat. Acad. Sci. U.S.A.* 92, 5111–5515. doi: 10.1073/pnas.92.11.5111

Ashburner, J., and Friston, K. J. (2000). Voxel-based morphometry-the methods. *Neuroimage.* 11, 805–821. doi: 10.1006/nimg.2000.0582

Aspell, J. E., Heydrich, L., Marillier, G., Lavanchy, T., Herbelin, B., Blanke, O., et al. (2013). Turning body and self-inside out: Visualized heartbeats alter bodily self-consciousness and tactile perception. *Psychol. Sci.* 24, 2445–2453. doi: 10.1177/0956797613498395

Baer, R. A., Smith, G. T., and Allen, K. B. (2004). Assessment of mindfulness by self-report: the Kentucky inventory of mindfulness skills. *Assessment* 11, 191–206. doi: 10.1177/1073191104268029

Baer, R. A., Smith, G. T., Hopkins, J., Krietemeyer, J., and Toney, L. (2006). Using self-report assessment methods to explore facets of mindfulness. *Assessment* 13, 27–45. doi: 10.1177/1073191105283504

Baker, S. C., Rogers, R. D., Owen, A. M., Frith, C. D., Dolan, R. J., Frackowiak, R. S. J., et al. (1996). Neural systems engaged by planning: a PET study of the Tower of London task. *Neuropsychologia* 34, 515–526. doi: 10.1016/0028-3932(95)00133-6

Balconi, M., Fronda, G., Venturella, I., and Crivelli, D. (2017). Conscious, pre-conscious and unconscious mechanisms in emotional behaviour. Some applications to the mindfulness approach with wearable devices. *Appl. Sci.* 7, 1–14. doi: 10.3390/app7121280

ben Salem, M., and Karlin, N. J. (2022). *Dispositional Mindfulness and Positive Mindset in Emerging Adult College Students: The Mediating Role of Decentering.* Psychological Reports. 1–19. doi: 10.1177/00332941211061705

Bergomi, C., Tschacher, W., and Kupper, Z. (2013). The assessment of mindfulness with self-report measures: existing scales and open issues. *Mindfulness* 4, 191–202. doi: 10.1007/s12671-012-0110-9

Bishop, S. R., Lau, M., Shapiro, S., Carlson, L., Anderson, N. D., Carmody, J., et al. (2004). Mindfulness: a proposed operational definition. *Clin. Psychol. Sci. Pract.* 11, 230–241. doi: 10.1093/clipsy.bph077

Biswal, B. B. (2012). Resting state fMRI: a personal history. *Neuroimage* 62, 938–944. doi: 10.1016/j.neuroimage.2012.01.090

Brefczynski-Lewis, J. A., Lutz, A., Schaefer, H. S., Levinson, D. B., and Davidson, R. J. (2007). Neural correlates of attentional expertise in long-term meditation practitioners. *PNAS.* 104, 111483–111488. Available online at: www.pnas.org/cgi/content/full/

Brewer, J. A., Worhunsky, P. D., Gray, J. R., Tang, Y.-Y., Weber, J., and Kober, H. (2011). Meditation experience is associated with differences in default mode network activity and connectivity. *Proc. Natl. Acad. Sci. U.S.A.* 108, 20254–20259. doi: 10.1073/pnas.1112029108

Brown, K. W., Goodman, R. J., and Inzlicht, M. (2013). Dispositional mindfulness and the attenuation of neural responses to emotional stimuli. *Soc. Cogn. Affect. Neurosci.* 8, 93–99. doi: 10.1093/scan/nss004

Brown, K. W., and Ryan, R. M. (2003). The benefits of being present: mindfulness and its role in psychological well-being. *J. Personal. Soc. Psychol.* 84, 822–848. doi: 10.1037/0022-3514.84.4.822

Brown, K. W., Ryan, R. M., and Creswell, J. D. (2007). Psychological inquiry copyright C. *Psychol. Inq.* 18, 272–281. doi: 10.1080/10478400701703344

Brugger, P., and Graves, R. E. (1997). Testing vs. believing hypotheses: magical ideation in the judgement of contingencies. *Cogn. Neuropsychiatry* 2, 251–272. doi: 10.1080/135468097396270

Buckner, R. L., Andrews-Hanna, J. R., and Schacter, D. L. (2008). The brain's default network. *Ann. N. Y. Acad. Sci.* 1124, 1–38. doi: 10.1196/annals.1440.011

Buckner, R. L., and Carroll, D. C. (2007). Self-projection and the brain. *Trends Cogn. Sci.* 11, 49–57. doi: 10.1016/j.tics.2006.11.004

Buckner, R. L., Raichle, M. E., Miezin, F. M., and Petersen, S. E. (1996). Functional anatomic studies of memory retrieval for auditory words and visual pictures. *J. Neurosci.* 16, 6219–6235. doi: 10.1523/JNEUROSCI.16-19-06219.1996

Bud Craig, A. D. (2009). How do you feel - now? The anterior insula and human awareness. *Nat. Rev. Neurosci.* 10, 59–70. doi: 10.1038/nrn2555

Bzdok, D., Langner, R., Schilbach, L., Engemann, D. A., Laird, A. R., Fox, P. T., et al. (2013). Segregation of the human medial prefrontal cortex in social cognition. *Front. Human Neurosci.* 7, 232. doi: 10.3389/fnhum.2013.00232

Calhoun, V. D., Adali, T., McGinty, V. B., Pekar, J. J., Watson, T. D., and Pearlson, G. D. (2001). fMRI activation in a visual-perception task: network of areas detected using the general linear model and independent components analysis. *Neuroimage* 14, 1080–1088. doi: 10.1006/nimg.2001.0921

Chalmers, D. J. (2013). How can we construct a science of consciousness? *Ann. N. Y. Acad. Sci.* 1303, 25–35. doi: 10.1111/nyas.12166

Chiesa, A. (2013). The difficulty of defining mindfulness: current thought and critical issues. *Mindfulness* 4, 255–268. doi: 10.1007/s12671-012-0123-4

Choifer, A. (2018). A new understanding of the first-person and third-person perspectives. *Philos. Papers* 47, 333–371. doi: 10.1080/05568641.2018.1450160

Choi-Kain, L. W., John, M. E., and Gunderson, G. (2008). Reviews and overviews mentalization: ontogeny, assessment, and application in the treatment of borderline personality disorder. *Am. J. Psychiatry* 165, 1127–1139. doi: 10.1176/appi.ajp.2008.07081360

Christoff, K., Cosmelli, D., Legrand, D., and Thompson, E. (2011). Specifying the self for cognitive neuroscience. *Trends Cogn. Sci.* 15, 104–112. doi: 10.1016/j.tics.2011.01.001

Christoff, K., Gordon, A. M., Smallwood, J., Smith, R., and Schooler, J. W. (2009). Experience sampling during fMRI reveals default network and executive system contributions to mind wandering. *Proc. Nat. Acad. Sci. U.S.A.* 106, 8719–8724. doi: 10.1073/pnas.0900234106

Christoff, K., Irving, Z. C., Fox, K. C. R., Spreng, R. N., and Andrews-Hanna, J. R. (2016). "Mind-wandering as spontaneous thought: A dynamic framework," in *Nature Reviews Neuroscience, Vol. 17* (Nature Publishing Group), 718–731. doi: 10.1038/nrn.2016.113

Cole, M. W., Reynolds, J. R., Power, J. D., Repovs, G., Anticevic, A., and Braver, T. S. (2013). Multi-task connectivity reveals flexible hubs for adaptive task control. *Nat. Neurosci.* 16, 1348–1355. doi: 10.1038/nn.3470

Colzato, L. S., Ozturk, A., and Hommel, B. (2012). Meditate to create: the impact of focused-attention and open-monitoring training on convergent and divergent thinking. *Front. Psychol.* 3, 116. doi: 10.3389/fpsyg.2012.00116

Courtney, S. M., Ungerleider, L. G., Keil, K., and Haxby, J., v. (1996). Object and spatial visual working memory activate separate neural systems in human cortex. *Cereb. Cortex* 6, 39–49. doi: 10.1093/cercor/6.1.39

Craig, A. D. (2002). How do you feel? Interoception: the sense of the physiological condition of the body. *Nat. Rev. Neurosci.* 3, 655–666. doi: 10.1038/nrn894

Craig, A. D. (2003). Interoception: The sense of the physiological condition of the body. *Curr. Opin. Neurobiol.* 13, 500–505. doi: 10.1016/S0959-4388(03)00090-4

Craig, A. D. (2004). Human feelings: why are some more aware than others? *Trends Cogn. Sci.* 8, 239–241. doi: 10.1016/j.tics.2004.04.004

Craig, A. D. (2009). Emotional moments across time: a possible neural basis for time perception in the anterior insula. *Philos. Trans. R. Soc. B Biol. Sci.* 364, 1933–1942. doi: 10.1098/rstb.2009.0008

Crane, N. T., Hayes, J. M., Viviano, R. P., Bogg, T., and Damoiseaux, J. S. (2020). Resting-state functional brain connectivity in a predominantly African-American sample of older adults: exploring links among personality traits, cognitive performance, and the default mode network. *Personal. Neurosci.* 3, 1–8. doi: 10.1017/pen.2020.4

Crane, R. S., and Hecht, F. M. (2018). Intervention integrity in mindfulness-based research. *Mindfulness* 9, 1370–1380. doi: 10.1007/s12671-018-0886-3

Culham, J. C., and Kanwisher, N. G. (2001). Culham_Cognitive Functions in human parietal cortex_2001. *Curr. Opin. Neurobiol.* 11, 157–163. doi: 10.1016/S0959-4388(00)00191-4

Dahl, C. J., Lutz, A., and Davidson, R. J. (2015). Reconstructing and deconstructing the self: Cognitive mechanisms in meditation practice. *Trends Cogn. Sci.* 19, 515–523. doi: 10.1016/j.tics.2015.07.001

Damasio, A. (1999). *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. New York, NY: Harcourt Brace and Company.

Dambrun, M., and Ricard, M. (2011). Self-centeredness and selflessness: a theory of self-based psychological functioning and its consequences for happiness. *Rev. General Psychol.* 15, 138–157. doi: 10.1037/a0023059

D'Argembeau, A. (2013). On the role of the ventromedial prefrontal cortex in self-processing: the valuation hypothesis. *Front. Human Neurosci.* 7, 372. doi: 10.3389/fnhum.2013.00372

Davenport, C., and Pagnini, F. (2016). Mindful learning: a case study of langerian mindfulness in schools. *Front. Psychol.* 7, 1372. doi: 10.3389/fpsyg.2016.01372

de Haan, S. (2021). Bio-psycho-social interaction: an enactive perspective. *Int. Rev. Psychiatry.* 33, 471–477. doi: 10.1080/09540261.2020.1830753

Debarnot, U., Sperduti, M., di Rienzo, F., and Guillot, A. (2014). Experts bodies, experts minds: how physical and mental training shape the brain. *Front. Human Neurosci.* 8, 280. doi: 10.3389/fnhum.2014.00280

Duncan, L. E., and Peterson, B. E. (2014). Authoritarianism, cognitive rigidity, and the processing of ambiguous visual information. *J. Soc. Psychol.* 154, 480–490. doi: 10.1080/00224545.2014.933764

Epstein, S. (2003). "Cognitive-experiential self-theory of personality," in *Comprehensive Handbook of Psychology, Vol. 5: Personality and Social Psychology, 1st Edn*., eds Milton. T and M. J. Lerner (Hoboken, NYL Wiley and Sons), 159–184.

Farb, N. A. S., Segal, Z., v, Mayberg, H., Bean, J., McKeon, D., Fatima, Z., et al. (2007). Attending to the present: mindfulness meditation reveals distinct neural modes of self-reference. *Soc. Cogn. Affect. Neurosci.* 2, 313–322. doi: 10.1093/scan/nsm030

Farb, N. A. S., Segal, Z. V., and Anderson, A. K. (2013). Mindfulness meditation training alters cortical representations of interoceptive attention. *Soc. Cogn. Affect. Neurosci.* 8, 15–26. doi: 10.1093/scan/nss066

Fivush, R., Habermas, T., Waters, T. E. A., and Zaman, W. (2011). The making of autobiographical memory: intersections of culture, narratives and identity. *Int. J. Psychol.* 46, 321–345. doi: 10.1080/00207594.2011.596541

Fox, K. C. R., Dixon, M. L., Nijeboer, S., Girn, M., Floman, J. L., Lifshitz, M., et al. (2016). Functional neuroanatomy of meditation: a review and meta-analysis of 78 functional neuroimaging investigations. *Neurosci. Biobehav. Rev.* 65, 208–228. doi: 10.1016/j.neubiorev.2016.03.021

Fox, K. C. R., Zakarauskas, P., Dixon, M., Ellamil, M., Thompson, E., and Christoff, K. (2012). Meditation experience predicts introspective accuracy. *PLoS ONE* 7, e45370. doi: 10.1371/journal.pone.0045370

Friston, K. J., Holmes, A. P., Poline, J.-B., Grasby, P. J., Williams, S. C. R., Frackowiak, R. S. J., et al. (1995). Analysis of fMRI time-series revisited. *Neuroimage* 2., 45–53 doi: 10.1006/nimg.1995.1007

Frith, U., and Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philos. Trans. R. Soc. B Biol. Sci.* 358, 459–473. doi: 10.1098/rstb.2002.1218

Fujino, M., Ueda, Y., Mizuhara, H., Saiki, J., and Nomura, M. (2018). Open monitoring meditation reduces the involvement of brain regions related to memory function. *Sci. Rep.* 8, 9968. doi: 10.1038/s41598-018-28274-4

Galijašević, M., Steiger, R., Regodić, M., Waibel, M., Sommer, P. J. D., Grams, A. E., et al. (2021). Brain energy metabolism in two states of mind measured by phosphorous magnetic resonance spectroscopy. *Front. Hum. Neurosci.* 15, 686433. doi: 10.3389/fnhum.2021.686433

Gallagher, S. (1997). Mutual enlightenment: recent phenomenology in cognitive science. *J. Consciousness Stud*. 146, 195–214.

Gallagher, S. (2000). Philosophical conceptions of the self: implications for cognitive science. *Trends Cogn. Neurosci.* 4, 14–21. doi: 10.1016/S1364-6613(99)01417-5

Gallagher, S., and Brøsted Sørensen, J. (2006). Experimenting with phenomenology. *Conscious. Cogn.* 15, 119–134. doi: 10.1016/j.concog.2005.03.002

Gallagher, S., and Varela, F. (2003). Redrawing the map and resetting the time: phenomenology and the cognitive sciences. *Can. J. Philos.* 29, 93–132. doi: 10.1080/00455091.2003.10717596

Garrison, K. A., Santoyo, J. F., Davis, J. H., Thornhill, I. V., T. A., Kerr, C. E., et al. (2013). Effortless awareness: using real time neurofeedback to investigate correlates of posterior cingulate cortex activity in meditators' self-report. *Front. Human Neurosci.* 7, 440. doi: 10.3389/fnhum.2013.00440

Garrison, K. A., Zeffiro, T. A., Scheinost, D., Constable, R. T., and Brewer, J. A. (2015). Meditation leads to reduced default mode network activity beyond an active task. *Cogn. Affect. Behav. Neurosci.* 15, 712–720. doi: 10.3758/s13415-015-0358-3

Gecht, J., Kessel, R., Forkmann, T., Gauggel, S., Drueke, B., Scherer, A., et al. (2014). A mediation model of mindfulness and decentering: sequential psychological constructs or one and the same? *BMC Psychol.* 2, 2–13. doi: 10.1186/2050-7283-2-18

Gethin, R. (2011). On some definitions of mindfulness. *Contemp. Buddhism* 12, 263–279. doi: 10.1080/14639947.2011.564843

Ghatan, P. H., Hsieh, J.-C., Wirsén-Meurling, A., Wredling, R., Eriksson, L., Stone-Elander, S., et al. (1995). Brain activation induced by the perceptual maze test: a PET study of cognitive performance. *Neuroimage* 2, 112–124. doi: 10.1006/nimg.1995.1014

Gianaros, P. J., Sheu, L. K., Matthews, K. A., Jennings, J. R., Manuck, S. B., and Hariri, A. R. (2008). Individual differences in stressor-evoked blood pressure reactivity vary with activation, volume, and functional connectivity of the amygdala. *J. Neurosci.* 28, 990–999. doi: 10.1523/JNEUROSCI.3606-07.2008

Goodman, M. J., and Schorling, J. B. (2012). A mindfulness course decreases burnout and improves well-being among healthcare providers. *Int. J. Psychiatry Med.* 43, 119–128. doi: 10.2190/PM.43.2.b

Greenberg, J., Reiner, K., and Meiran, N. (2010). "Mind the trap": mindfulness practice reduces cognitive rigidity. *PLoS ONE* 5, e36206. doi: 10.1371/journal.pone.0036206

Gross, J. J. (2001). Emotion regulation in adulthood: timing is everything. *Direct. Psychol. Sci.* 10, 214–219. doi: 10.1111/1467-8721.00152

Gusnard, D. A., Akbudak, E., Shulman, G. L., and Raichle, M. E. (2001). Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. *PNAS* 98, 4259–4264. doi: 10.1073/pnas.071043098

Gusnard, D. A., and Raichle, M. E. (2001). CBF Baseline_Raichle_2001. *Nat. Rev. Neurosci.* 2, 685–694. doi: 10.1038/35094500

Harris, S., Sheth, S. A., and Cohen, M. S. (2008). Functional neuroimaging of belief, disbelief, and uncertainty. *Ann. Neurol.* 63, 141–147. doi: 10.1002/ana.21301

Hasenkamp, W., and Barsalou, L. W. (2012). Effects of meditation experience on functional connectivity of distributed brain networks. *Front. Hum. Neurosci.* 6, 38. doi: 10.3389/fnhum.2012.00038

Hasenkamp, W., Wilson-Mendenhall, C. D., Duncan, E., and Barsalou, L. W. (2012). Mind wandering and attention during focused meditation: a fine-grained temporal analysis of fluctuating cognitive states. *Neuroimage* 59, 750–760. doi: 10.1016/j.neuroimage.2011.07.008

Haxby, J., v, Horwitz, B., Ungerleider, L. G., Maisog, J. M., Pietrini, P., and Grady, C. L. (1994). The functional organization of human extrastriate cortex: a PET-rCBF study of selective attention to faces and locations. *J. Neurosci.* 74, 6338–6353. doi: 10.1523/JNEUROSCI.14-11-06336.1994

Hernández, S. E., Barros-Loscertales, A., Xiao, Y., González-Mora, J. L., and Rubia, K. (2018). Gray matter and functional connectivity in anterior cingulate cortex are associated with the state of mental silence during Sahaja Yoga Meditation. *Neuroscience* 371, 395–406. doi: 10.1016/j.neuroscience.2017.12.017

Herwig, U., Kaffenberger, T., Jäncke, L., and Brühl, A. B. (2010). Self-related awareness and emotion regulation. *Neuroimage* 50, 734–741. doi: 10.1016/j.neuroimage.2009.12.089

Herwig, U., Opialla, S., Cattapan, K., Wetter, T. C., Jäncke, L., and Brühl, A. B. (2018). Emotion introspection and regulation in depression. *Psychiatry Res. Neuroimaging* 277, 7–13. doi: 10.1016/j.pscychresns.2018.04.008

Hölzel, B. K., Lazar, S. W., Gard, T., Schuman-Olivier, Z., Vago, D. R., and Ott, U. (2011). How does mindfulness meditation work? Proposing mechanisms of action from a conceptual and neural perspective. *Perspect. Psychol. Sci.* 6, 537–559. doi: 10.1177/1745691611419671

Hölzel, B. K., Ott, U., Gard, T., Hempel, H., Weygandt, M., Morgen, K., et al. (2008). Investigation of mindfulness meditation practitioners with voxel-based morphometry. *Soc. Cogn. Affect. Neurosci.* 3, 55–61. doi: 10.1093/scan/nsm038

Howlett, J. R., and Paulus, M. P. (2015). The neural basis of testable and non-testable beliefs. *PLoS ONE* 10, e124596. doi: 10.1371/journal.pone.0124596

Izmirli, I. M. (2014). Wittengstein's language games and forms of life from a social constructivist point of view. *Am. J. Educ. Res.* 2, 291–298. doi: 10.12691/education-2-5-9

James, W. (1890). *The Principles of Psychology*, Vol. 1. Henry Holt and Company. Available online at: https://archive.org/details/theprinciplesofp01jameuoft doi: 10.1037/10538-000

Jensen, C. G., Vangkilde, S., Frokjaer, V., and Hasselbalch, S. G. (2012). Mindfulness training affects attention-or is it attentional effort? *J. Exp. Psychol. General* 141, 106–123. doi: 10.1037/a0024931

Jha, A. P., Stanley, E. A., Kiyonaga, A., Wong, L., and Gelfand, L. (2010). Examining the protective effects of mindfulness training on working memory capacity and affective experience. *Emotion* 10, 54–64. doi: 10.1037/a0018438

Josipovic, Z. (2014). Neural correlates of nondual awareness in meditation. *Ann. N. Y. Acad. Sci.* 1307, 9–18. doi: 10.1111/nyas.12261

Josipovic, Z., Dinstein, I., Weber, J., and Heeger, D. J. (2011). Influence of meditation on anti-correlated networks in the brain. *Front. Hum. Neurosci.* 5, 183. doi: 10.3389/fnhum.2011.00183

Kabat-Zinn, J. (1990). *Full Catastrophe Living: Using the Wisdom of Your Mind to Face Stress, Pain and Illness*. New York, NY: Delacorte.

Kabat-Zinn, J., Lipworth, L., and Burney, R. (1985). The clinical use of mindfulness meditation for the self-regulation of chronic pain. *J. Behav. Med.* 8, 163–190. doi: 10.1007/BF00845519

Kawashima, R., Roland, P. E., and O'Sullivan, B. T. (1994). Fields in human motor areas involved in preparation for reaching, actual reaching, and visuomotor learning: a positron emission tomography study. *J. Neurosci.* 14, 3462–3474. doi: 10.1523/JNEUROSCI.14-06-03462.1994

Keye, M. D., and Pidgeon, A. M. (2013). Investigation of the relationship between resilience, mindfulness, and academic self-efficacy. *Open J. Soc. Sci.* 1, 1–4. doi: 10.4236/jss.2013.16001

Killingsworth, M. A., and Gilbert, D. T. (2010). A wandering mind is an unhappy mind. *Science* 330, 932. doi: 10.1126/science.1192439

Kral, T. R. A., Schuyler, B. S., Mumford, J. A., Rosenkranz, M. A., Lutz, A., and Davidson, R. J. (2018). Impact of short- and long-term mindfulness meditation training on amygdala reactivity to emotional stimuli. *Neuroimage* 181, 301–313. doi: 10.1016/j.neuroimage.2018.07.013

Krammer, W., Missimer, J. H., Habegger, S., Pastore-Wapp, M., Wiest, R., and Weder, B. J. (2020). Sensing form - finger gaiting as key to tactile object exploration - a data glove analysis of a prototypical daily task. *J. Neuroeng. Rehabil.* 17, 1–17. doi: 10.1186/s12984-020-00755-6

Kurczek, J., Wechsler, E., Ahuja, S., Jensen, U., Cohen, N. J., Tranel, D., et al. (2015). Differential contributions of hippocampus and medial prefrontal cortex to self-projection and self-referential processing. *Neuropsychologia* 73, 116–126. doi: 10.1016/j.neuropsychologia.2015.05.002

Kwak, S., Lee, T. Y., Jung, W. H., Hur, J.-W., Bae, D., Hwang, W. J., et al. (2019). The immediate and sustained positive effects of meditation on resilience are mediated by changes in the resting brain. *Front. Hum. Neurosci.* 13, 101. doi: 10.3389/fnhum.2019.00101

Kyselo, M. (2014). The body social: an enactive approach to the self. *Front. Psychol.* 5, 986. doi: 10.3389/fpsyg.2014.00986

Langer, E., Djikic, M., Pirson, M., Madenci, A., and Donohue, R. (2010). Believing is seeing: using mindlessness (mindfully) to improve visual acuiy. *Psychol. Sci.* 21, 661–666. doi: 10.1177/0956797610366543

Langer, E. J. (2014). *Mindfulness, 2nd Edn.* Boston, MA: Da Capo Press.

Langer, E. J., and Moldoveanu, M. (2000). The construct of mindfulness. *J. Soc. Issues* 56, 1–9. doi: 10.1111/0022-4537.00148

Lau, M. A., Bishop, S. R., Segal, Z. v., Buis, T., Anderson, N. D., Carlson, L., et al. (2006). The Toronto mindfulness scale: development and validation. *J. Clin. Psychol.* 62, 1445–1467. doi: 10.1002/jclp.20326

Lazar, S. W., Kerr, C. E., Wasserman, R. H., Gray, J. R., Greve, D. N., Treadway, M. T., et al. (2005). Meditation experience is associated with increased cortical thickness. *Neuroreport* 16, 1893–1897. doi: 10.1097/01.wnr.0000186598.66243.19

Lebois, L. A. M., Papies, E. K., Gopinath, K., Cabanban, R., Quigley, K. S., Krishnamurthy, V., et al. (2015). A shift in perspective: decentering through mindful attention to imagined stressful events. *Neuropsychologia* 75, 505–524. doi: 10.1016/j.neuropsychologia.2015.05.030

Leech, R., Kamourieh, S., Beckmann, C. F., and Sharp, D. J. (2011). Fractionating the default mode network: distinct contributions of the ventral and dorsal posterior cingulate cortex to cognitive control. *J. Neurosci.* 31, 3217–3224. doi: 10.1523/JNEUROSCI.5626-10.2011

Legrand, D. (2007). Pre-reflective self-as-subject from experiential and empirical perspectives. *Conscious. Cogn.* 16, 583–599. doi: 10.1016/j.concog.2007.04.002

Lenhart, L., Steiger, R., Waibel, M., Mangesius, S., Grams, A. E., Singewald, N., et al. (2020). Cortical reorganization processes in meditation naïve participants induced by 7 weeks focused attention meditation training. *Behav. Brain Res.* 395, 112828. doi: 10.1016/j.bbr.2020.112828

Livesey, A. C., Wall, M. B., and Smith, A. T. (2007). Time perception: Manipulation of task difficulty dissociates clock functions from other cognitive demands. *Neuropsychologia* 45, 321–333. doi: 10.1016/j.neuropsychologia.2006.06.033

Logothetis, N. K., and Pfeuffer, J. (2004). On the nature of the BOLD fMRI contrast mechanism. *Magn. Reson. Imaging* 22, 1517–1531. doi: 10.1016/j.mri.2004.10.018

Lu, H., Song, Y., Xu, M., Wang, X., Li, X., and Liu, J. (2014). The brain structure correlates of individual differences in trait mindfulness: a voxel-based morphometry study. *Neuroscience* 272, 21–28. doi: 10.1016/j.neuroscience.2014.04.051

Luders, E., Kurth, F., Mayer, E. A., Toga, A. W., Narr, K. L., and Gaser, C. (2012). The unique brain anatomy of meditation practitioners: alterations in cortical gyrification. *Front. Hum. Neurosci.* 6, 34. doi: 10.3389/fnhum.2012.00034

Lutz, A., Jha, A. P., Dunne, J. D., and Saron, C. D. (2015). Investigating the phenomenological matrix of mindfulness-related practices from a neurocognitive perspective. *Am. Psychol.* 70, 632–658. doi: 10.1037/a0039585

Lutz, A., Slagter, H. A., Dunne, J. D., and Davidson, R. J. (2008). Attention regulation and monitoring in meditation. *Trends Cogn. Sci.* 12, 163–169. doi: 10.1016/j.tics.2008.01.005

Malagurski, B., Liem, F., Oschwald, J., Mérillat, S., and Jäncke, L. (2020). Functional dedifferentiation of associative resting state networks in older adults - a longitudinal study. *Neuroimage* 214, 4829–4845. doi: 10.1016/j.neuroimage.2020.116680

Mansouri, F. A., Tanaka, K., and Buckley, M. J. (2009). Conflict-induced behavioural adjustment: a clue to the executive functions of the prefrontal cortex. *Nat. Rev. Neurosci.* 10, 141–152. doi: 10.1038/nrn2538

Mason, M. F., Norton, M. I., van Horn, J. D., Wegner, D. M., Grafton, S. T., and Macrae, C. N. (2007a). Wandering minds: the default network and stimulus-independent thought. *Science* 315, 393–395. doi: 10.1126/science.1131295

Mason, M. F., Norton, M. I., van Horn, J. D., Wegner, D. M., Grafton, S. T., and Macrae, C. N. (2007b). Response to comment on "wandering minds: the default network and stimulus-independent thought". *Science* 317, 43. doi: 10.1126/science.1141078

Mazoyer, B., Zago, L., Mellet, E., Bricogne, S., Etard, O., Houdé, O., et al. (2001). Cortical networks for working memory and executive functions sustain the conscious resting state in man. *Brain Res. Bull.* 54, 287–298. doi: 10.1016/S0361-9230(00)00437-8

McEvoy, P. M., Mahoney, A. E. J., and Moulds, M. L. (2010). Are worry, rumination, and post-event processing one and the same? *J. Anxiety Disord.* 24, 509–519. doi: 10.1016/j.janxdis.2010.03.008

Metzinger, T. (2018). Why is mind wandering interesting for philosophers? *The Oxford Handbook of Sontaneous Thought: The Oxford Handbook of Sontaneous Thought: Mind-wandering, Creativity, Dreaming, and Clinical Conditions*, eds K. C. R. Fox and K. Christoff (Oxford: The Oxford University Press), 97–112.

Mikulas, W. L. (2011). Mindfulness: significant common confusions. *Mindfulness* 2, 1–7. doi: 10.1007/s12671-010-0036-z

Miller, E. K. (2000). The prefrontal cortex and cognitive control. *Nat. Rev. Neurosci.* 1, 59–65. doi: 10.1038/35036228

Modinos, G., Ormel, J., and Aleman, A. (2010). Individual differences in dispositional mindfulness and brain activity involved in reappraisal of emotion. *Soc. Cogn. Affect. Neurosci.* 5, 369–377. doi: 10.1093/scan/nsq006

Moore, A., and Malinowski, P. (2009). Meditation, mindfulness and cognitive flexibility. *Conscious. Cogn.* 18, 176–186. doi: 10.1016/j.concog.2008.12.008

Mrazek, M. D., Franklin, M. S., Phillips, D. T., Baird, B., and Schooler, J. W. (2013). Mindfulness training improves working memory capacity and GRE performance while reducing mind wandering. *Psychol. Sci.* 24, 776–781. doi: 10.1177/0956797612459659

Murakami, H., Katsunuma, R., Oba, K., Terasawa, Y., Motomura, Y., Mishima, K., et al. (2015). Neural networks for mindfulness and emotion suppression. *PLoS ONE* 10, e0128005. doi: 10.1371/journal.pone.0128005

Murakami, H., Nakao, T., Matsunaga, M., Kasuya, Y., Shinoda, J., Yamada, J., et al. (2012). The structure of mindful brain. *PLoS ONE* 7, e46377. doi: 10.1371/journal.pone.0046377

Neff, K. (2003). Self-compassion: an alternative conceptualization of a healthyattitudetoward oneself. *Self Identity* 2, 85–101. doi: 10.1080/15298860309032

Nehaniv, C. L., Forster, F., Saunders, J., Broz, F., Antonova, E., Kose, H., et al. (2013). *Interaction and Experience in Enactive Intelligence and Humanoid Robotics Artificial Life (ALife)*. (Singapore: IEEE Symposium), 148–155. doi: 10.1109/ALIFE.2013.6602445

Nolen-Hoeksema, S. (1991). Responses to depression and their effects on the duration of depressive episodes. *J. Abnorm. Psychol.* 100, 569–582. doi: 10.1037/0021-843X.100.4.569

Northoff, G., and Bermpohl, F. (2004). Cortical midline structures and the self. *Trends Cogn. Sci.* 8, 102–107. doi: 10.1016/j.tics.2004.01.004

Northoff, G., Heinzel, A., de Greck, M., Bermpohl, F., Dobrowolny, H., and Panksepp, J. (2006). Self-referential processing in our brain-a meta-analysis of imaging studies on the self. *Neuroimage* 31, 440–457. doi: 10.1016/j.neuroimage.2005.12.002

Olivares, F. A., Vargas, E., Fuentes, C., Martínez-Pernía, D., and Canales-Johnson, A. (2015). Neurophenomenology revisited: Second-person methods for the study of human consciousness. *Front. Psychol.* 6, 673. doi: 10.3389/fpsyg.2015.00673

Opialla, S., Lutz, J., Scherpiet, S., Hittmeyer, A., Jäncke, L., Rufer, M., et al. (2015). Neural circuits of emotion regulation: a comparison of mindfulness-based and cognitive reappraisal strategies. *Eur. Arch. Psychiatry Clin. Neurosci.* 265, 45–55. doi: 10.1007/s00406-014-0510-z

Pasquali, A., Timmermans, B., and Cleeremans, A. (2010). Know thyself: metacognitive networks and measures of consciousness. *Cognition* 117, 182–190. doi: 10.1016/j.cognition.2010.08.010

Pauen, M. (2012). The second-person perspective. *Inquiry* 55, 33–49. doi: 10.1080/0020174X.2012.643623

Posner, M. I. (1980). Orienting of attention*. *Q. J. EJrperimental Psychol.* 32, 3–25. doi: 10.1080/00335558008248231

Preckel, K., Kanske, P., and Singer, T. (2018). On the interaction of social affect and cognition: empathy, compassion and theory of mind. *Curr. Opin. Behav. Sci.* 19, 1–6. doi: 10.1016/j.cobeha.2017.07.010

Raffone, A., Marzetti, L., del Gratta, C., Perrucci, M. G., Romani, G. L., and Pizzella, V. (2019). Toward a brain theory of meditation. *Progress Brain Res.* 244, 207–232. doi: 10.1016/bs.pbr.2018.10.028

Revel, J.-F., and Ricard, M. (1999). *The Monk and the Philosopher: A Father and Son Discuss the Meaning of Life*. New York, NY: Schocken Books.

Rimes, K. A., and Wingrove, J. (2011). Pilot study of mindfulness-based cognitive therapy for trainee clinical psychologists. *Behav. Cogn. Psychother.* 39, 235–241. doi: 10.1017/S1352465810000731

Roeser, R. W., and Peck, S. C. (2009). An education in awareness: self, motivation, and self-regulated learning in contemplative perspective. *Educ. Psychol.* 44, 119–136. doi: 10.1080/00461520902832376

Rummel, J., and Boywitt, C. D. (2014). Controlling the stream of thought: working memory capacity predicts adjustment of mind-wandering to situational demands. *Psychon. Bull. Rev.* 21, 1309–1315. doi: 10.3758/s13423-013-0580-3

Sacks, O., and Hirsch, J. (2008). A neurology of belief. *Ann. Neurol.* 63, 129–130. doi: 10.1002/ana.21378

Santarnecchi, E., Egiziano, E., D'Arista, S., Gardi, C., Romanella, S. M., Mencarelli, L., et al. (2021). Mindfulness-based stress reduction training modulates striatal and cerebellar connectivity. *J. Neurosci. Res.* 99, 1236–1252. doi: 10.1002/jnr.24798

Scheibner, H. J., Bogler, C., Gleich, T., Haynes, J. D., and Bermpohl, F. (2017). Internal and external attention and the default mode network. *Neuroimage* 148, 381–389. doi: 10.1016/j.neuroimage.2017.01.044

Schurz, M., Aichhorn, M., Martin, A., and Perner, J. (2013). Common brain areas engaged in false belief reasoning and visual perspective taking: a meta-analysis of functional brain imaging studies. *Front. Human Neurosci.* 7, 712. doi: 10.3389/fnhum.2013.00712

Searle, J. R. (1994). *The Rediscovery of the Mind, 1st Edn*. MIT Press. Available online at: https://books.google.ch/books?hl=de&lr=&id=eoh8e52wo_oC&oi=fnd&pg=PR9&dq=Searle+John+R+The+Rediscovery+of+the+Mind+and&ots=tCQQArXmUR&sig=eLD_hrxvd1p5R1kwwW-O--pxulc#v=onepage&q=Searle%20John%20R%20The%20Rediscovery%20of%20the%20Mind%20and&f=false

Seitz, R. J., Angel, H.-F., and Paloutzian, R. F. (2022). Statements of believing involve attribution. *Academia Lett.* 4624. doi: 10.20935/AL4624

Seitz, R. J., Paloutzian, R. F., and Angel, H. F. (2016). Processes of believing: Where do they come from? What are they good for? *F1000Research* 5, 1–20. doi: 10.12688/f1000research.9773.1

Seitz, R. J., Paloutzian, R. F., and Angel, H. F. (2018). From believing to belief: a general theoretical model. *J. Cogn. Neurosci.* 30, 1254–1264. doi: 10.1162/jocn_a_01292

Shackman, A. J., Salomons, T. v., Slagter, H. A., Fox, A. S., Winter, J. J., and Davidson, R. J. (2011). The integration of negative affect, pain and cognitive control in the cingulate cortex. *Nat. Rev. Neurosci.* 12, 154–167. doi: 10.1038/nrn2994

Shapiro, S. L., Carlson, L. E., Astin, J. A., and Freedman, B. (2006). Mechanisms of mindfulness. *J. Clin. Psychol.* 62, 373–386. doi: 10.1002/jclp.20237

Shapiro, S. L., and Schwartz, G. E. (2000). Intentional systemic mindfulness: an integrative model for self-regulation and health. *Adv. Mind Body Med.* 16, 128–134. doi: 10.1054/ambm.1999.0118

Shaurya Prakash, R., de Leon, A. A., Klatt, M., Malarkey, W., and Patterson, B. (2013). Mindfulness disposition and default-mode network connectivity in older adults. *Soc. Cogn. Affect. Neurosci.* 8, 112–117. doi: 10.1093/scan/nss115

Shepherd, K. A., Coifman, K. G., Matt, L. M., and Fresco, D. M. (2016). Development of a self-distancing task and initial validation of responses. *Psychol. Assess.* 28, 841–855. doi: 10.1037/pas0000297

Sheppes, G., and Gross, J. J. (2011). Is timing everything? Temporal considerations in emotion regulation. *Personal. Soc. Psychol. Rev.* 15, 319–331. doi: 10.1177/1088868310395778

Shiota, M. N., and Levenson, R. W. (2009). Effects of aging on experimentally instructed detached reappraisal, positive reappraisal, and emotional behavior suppression. *Psychol. Aging* 24, 890–900. doi: 10.1037/a0017896

Shulman, G. L., Fiez, J. A., Corbetta, M., Buckner, R. L., Miezin, F. M., Raichle, M. E., et al. (1997). Common blood flow changes across visual tasks: II.: decreases in cerebral cortex. *J. Cogn. Neurosci.* 9, 648–663. doi: 10.1162/jocn.1997.9.5.648

Singer, J. L., and Antrobus, J. S. (1970). *Imaginal Process Inventory*. Available online at: http://neuroinformatics.harvard.edu/w/public/images/5/55/Ipi.pdf

Slagter, H. A., Lutz, A., Greischar, L. L., Francis, A. D., Nieuwenhuis, S., Davis, J. M., et al. (2007). Mental training affects distribution of limited brain resources. *PLoS Biol.* 5, e138. doi: 10.1371/journal.pbio.0050138

Smallwood, J., Fitzgerald, A., Miles, L. K., and Phillips, L. H. (2009). Shifting moods, wandering minds: negative moods lead the mind to wander. *Emotion* 9, 271–276. doi: 10.1037/a0014855

Smallwood, J., and O'Connor, R. C. (2011). Imprisoned by the past: unhappy moods lead to a retrospective bias to mind wandering. *Cogn. Emot.* 25, 1481–1490. doi: 10.1080/02699931.2010.545263

Smoski, M. J., Keng, S.-L., Ji, J. L., Moore, T., Minkel, J., and Dichter, G. S. (2015). Neural indicators of emotion regulation via acceptance vs reappraisal in remitted major depressive disorder. *Soc. Cogn. Affect. Neurosci.* 10, 1187–1194. doi: 10.1093/scan/nsv003

Stawarczyk, D., Majerus, S., and D'Argembeau, A. (2013). Concern-induced negative affect is associated with the occurrence and content of mind-wandering. *Conscious. Cogn.* 22, 442–448. doi: 10.1016/j.concog.2013.01.012

Stephan, E., Pardo, J. v., Faris, P. L., Hartman, B. K., Kim, S. W., Ivanov, E. H., et al. (2003). Functional neuroimaging of gastric distention. *J. Gastrointest. Surg.* 7, 740–749. doi: 10.1016/S1091-255X(03)00071-4

Sterzer, P., and Kleinschmidt, A. (2010). Anterior insula activations in perceptual paradigms: often observed but barely understood. *Brain Struct. Funct.* 214, 611–622. doi: 10.1007/s00429-010-0252-2

Takano, K., and Tanno, Y. (2009). Self-rumination, self-reflection, and depression: self-rumination counteracts the adaptive effect of self-reflection. *Behav. Res. Ther.* 47, 260–264. doi: 10.1016/j.brat.2008.12.008

Tang, Y.-Y., Ma, Y., Wang, J., Fan, Y., Feng, S., Lu, Q., et al. (2007). Short-term meditation training improves attention and self-regulation. *Proc. Natl. Acad. Sci. U.S.A.* 104, 17152–17156. doi: 10.1073/pnas.0707678104

Tang, Y. Y., Hölzel, B. K., and Posner, M. I. (2015). The neuroscience of mindfulness meditation. *Nat. Rev. Neurosci.* 16, 213–225. doi: 10.1038/nrn3916

Taren, A. A., Gianaros, P. J., Greco, C. M., Lindsay, E. K., Fairgrieve, A., Brown, K. W., et al. (2014). Mindfulness meditation training alters stress-related amygdala resting state functional connectivity: a randomized controlled trial. *Soc. Cogn. Affect. Neurosci.* 10, 1758–1768. doi: 10.1093/scan/nsv066

Taylor, V. A., Daneault, V., Grant, J., Scavone, G., Breton, E., Roffe-Vidal, S., et al. (2013). Impact of meditation training on the default mode network during a restful state. *Soc. Cogn. Affect. Neurosci.* 8, 4–14. doi: 10.1093/scan/nsr087

Teasdale, J. D., Moore, R. G., Hayhurst, H., Pope, M., Williams, S., Segal, Z., et al. (2002). Metacognitive awareness and prevention of relapse in depression:

empirical evidence. *J. Consult. Clin. Psychol.* 70, 275–287. doi: 10.1037/0022-006X.70.2.275

Teasdale, J. D., Segal, Z., v, Williams, J. M., Ridgeway, V. A., Soulsby, J. M., and Lau, M. A. (2000). Prevention of relapse/recurrence in major depression by mindfulness-based cognitive therapy. *J. Consult. Clin. Psychol.* 68, 615–623. doi: 10.1037/0022-006X.68.4.615

Todd, R. M., Cunningham, W. A., Anderson, A. K., and Thompson, E. (2012). Affect-biased attention as emotion regulation. *Trends Cogn. Sci.* 16, 365–372. doi: 10.1016/j.tics.2012.06.003

Trapnell, P. D., and Campbell, J. D. (1999). Private self-consciousness and the five factor model of personality: distinguishing rumination from reflection. *J. Pers. Soc. Psychol.* 76, 284–304. doi: 10.1037/0022-3514.76.2.284

Travis, F. (2014). Transcendental experiences during meditation practice. *Ann. N. Y. Acad. Sci.* 1307, 1–8. doi: 10.1111/nyas.12316

Vago, D. R. (2014). Mapping modalities of self-awareness in mindfulness practice: a potential mechanism for clarifying habits of mind. *Ann. N. Y. Acad. Sci.* 1307, 28–42. doi: 10.1111/nyas.12270

Vago, D. R., and Silbersweig, D. A. (2012). Self-awareness, self-regulation, and self-transcendence (S-ART): a framework for understanding the neurobiological mechanisms of mindfulness. *Front. Hum. Neurosci.* 6, 296. doi: 10.3389/fnhum.2012.00296

Vago, D. R., and Zeidan, F. (2016). The brain on silent: mind wandering, mindful awareness, and states of mental tranquility. *Ann. N. Y. Acad. Sci.* 1373, 96–113. doi: 10.1111/nyas.13171

Valentine, E. R., and Sweet, P. L. G. (1999). Meditation and attention: a comparison of the effects of concentrative and mindfulness meditation on sustained attention. *Ment. Health Relig. Cult.* 2, 56–70. doi: 10.1080/13674679908406332

van Buuren, M., Gladwin, T. E., Zandbelt, B. B., Kahn, R. S., and Vink, M. (2010). Reduced functional coupling in the default-mode network during self-referential processing. *Hum. Brain Mapp.* 31, 1117–1127. doi: 10.1002/hbm.20920

van Vugt, M. K., and Jha, A. P. (2011). Investigating the impact of mindfulness meditation training on working memory: a mathematical modeling approach. *Cogn. Affect. Behav. Neurosci.* 11, 344–353. doi: 10.3758/s13415-011-0048-8

Vann, S. D., Aggleton, J. P., and Maguire, E. A. (2009). What does the retrosplenial cortex do? *Nat. Rev. Neurosci.* 10, 792–802. doi: 10.1038/nrn2733

Varela, F. (1996). Neurophenomenology: a methodological remedy to the hard problem. *J. Consciousness Stud.* 3, 330–349.

Varela, F., and Shear, J. (1999). Mendeley reference manager. *J. Consciousness Stud.* 6, 1–14.

Varela, F. J., Thompson, E., and Rosch, E. (2016). *The Embodied Mind: Cognitive Science and Human Experience. Revised Edition*. Cambridge, MA: MIT Press.

Verfaellie, M., Wank, A. A., Reid, A. G., Race, E., and Keane, M. M. (2019). Self-related processing and future thinking: distinct contributions of ventromedial prefrontal cortex and the medial temporal lobes. *Cortex* 115, 159–171. doi: 10.1016/j.cortex.2019.01.028

Vidotto, G., Massidda, D., Noventa, S., and Vicentini, M. (2012). Trusting beliefs: a functional measurement study. *Psicologica.* 33, 575–590. Available online at: https://files.eric.ed.gov/fulltext/EJ980495.pdf

Vincent, J. L., Kahn, I., Snyder, A. Z., Raichle, M. E., and Buckner, R. L. (2008). Evidence for a frontoparietal control system revealed by intrinsic functional connectivity. *J. Neurophysiol.* 100, 3328–3342. doi: 10.1152/jn.90355.2008

Vogt, B. A., Vogt, L., and Laureys, S. (2006). Cytology and functionally correlated circuits of human posterior cingulate areas. *Neuroimage* 29, 452–466. doi: 10.1016/j.neuroimage.2005.07.048

Walach, H., Buchheld, N., Buttenmüller, V., Kleinknecht, N., and Schmidt, S. (2006). Measuring mindfulness-the Freiburg Mindfulness Inventory (FMI). *Pers. Individ. Dif.* 40, 1543–1555. doi: 10.1016/j.paid.2005.11.025

Wang, X., Wu, Q., Egan, L., Gu, X., Liu, P., Gu, H., et al. (2019). Anterior insular cortex plays a critical role in interoceptive attention. *ELife* 8, 1–31. doi: 10.7554/eLife.42265

Wang, X., Xu, M., Song, Y., Li, X., Zhen, Z., Yanga, Z., et al. (2014). The network property of the thalamus in the default mode network is correlated with trait mindfulness. *Neuroscience* 278, 291–301. doi: 10.1016/j.neuroscience.2014.08.006

Wilkins, A. J., Shallice, T., and McCarthy, R. (1987). Frontal lesions and sustained attention. *Neuropsychologia* 25, 359–365. doi: 10.1016/0028-3932(87)90024-8

Wittmann, M., and Schmidt, S. (2014). "Mindfulness meditation and the experience of time," in *Meditation - Neuroscientific Approaches and Philosophical Implications, Studies in Neuroscience, Consciousness and Spirituality 2*, eds S. Schmidt and H. Walach (Cham: Springer International Publishing), 199–210.

Wozniak, M. (2018). "I" and "Me": the self in the context of consciousness. *Front. Psychol.* 9, 1656 doi: 10.3389/fpsyg.2018.01656

Zhuang, K., Bi, M., Li, Y., Xia, Y., Guo, X., Chen, Q., et al. (2017). A distinction between two instruments measuring dispositional mindfulness and the correlations between those measurements and the neuroanatomical structure. *Sci. Rep.* 7, 6252. doi: 10.1038/s41598-017-06599-w

frontiers | Frontiers in Psychology

TYPE Opinion
PUBLISHED 28 September 2022
DOI 10.3389/fpsyg.2022.943255

Check for updates

OPEN ACCESS

EDITED BY
Rüdiger J. Seitz,
Heinrich Heine University of Düsseldorf, Germany

REVIEWED BY
Gao Pan,
First Affiliated Hospital of Anhui Medical University, China

*CORRESPONDENCE
Aljoscha C. Neubauer
aljoscha.neubauer@uni-graz.at

SPECIALTY SECTION
This article was submitted to Cognition, a section of the journal Frontiers in Psychology

RECEIVED 13 May 2022
ACCEPTED 07 September 2022
PUBLISHED 28 September 2022

CITATION
Neubauer AC and Hofer G (2022) Believing in one's abilities: Ability estimates as a form of beliefs. *Front. Psychol.* 13:943255. doi: 10.3389/fpsyg.2022.943255

COPYRIGHT
© 2022 Neubauer and Hofer. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Believing in one's abilities: Ability estimates as a form of beliefs

Aljoscha C. Neubauer* and Gabriela Hofer

Section of Differential Psychology, Institute of Psychology, University of Graz, Graz, Austria

KEYWORDS

intelligence, abilities, self-estimate, other-estimate, beliefs, credition

Introduction

What people think of their own abilities (e.g., whether they see themselves as particularly intelligent, creative, or emotionally competent) has been the topic of a lot of psychological research. In a recent book chapter (Neubauer and Hofer, 2020), we provided a detailed review of this topic. Here, we highlight parallels between estimates of (or beliefs in) one's abilities and work on broader beliefs and the process of believing (also termed "credition"; Angel, 2013).

Abilities and what people know about theirs

Psychological concepts like abilities, skills, competencies, and talents have a long tradition in differential psychology (i.e., the study of individual differences in human psychological traits; Cooper, 2020). The "via regia" to assessing these traits in research and applied settings (e.g., human resources) are psychometric ability tests, like tests of intelligence, social skills, creative potentials, or attention. People's scores in such tests predict important outcomes such as professional success (e.g., Schmidt and Hunter, 2004; Harari et al., 2016) or well-being (e.g., Acar et al., 2021). However, these tests are (1) challenging to develop and (2) often costly and time-consuming to administer.

Around 100 years ago (e.g., Cogan et al., 1915), the idea of potentially more economic proxies of abilities came up: People could simply estimate their abilities in a given domain (e.g., verbal, numerical, or visuospatial abilities; for a review, see Neubauer and Hofer, 2020) by reporting their agreement to statements (e.g., Neubauer et al., 2018) like

- "I can easily rephrase a text using different wording."
- "I have good mental arithmetic skills." or
- "I am good at finding my way in an unknown area."

But self-estimates are not only used in standardized psychological assessments: People also assess their own abilities in everyday situations, for example, before taking an exam, when deciding on a career, or even before crossing a street (see also Ackerman and Wolman, 2007; Neubauer and Hofer, 2020). Thus, self-estimates can guide behavior (e.g., Ackerman and Wolman, 2007). They also show considerable overlap with other well-researched psychological constructs, such as self-esteem (Rosenberg, 1965), self-efficacy (Bandura, 1977), or self-concept (e.g., Marsh, 1990), all of which tapping into the positivity of people's self-views (see also Ackerman and Wolman, 2007; Marsh et al., 2019).

The pervasiveness of self-estimates leads us to an important question: How well do these subjective judgments correspond to objective performance assessments? In the last 100 years, dozens of empirical studies tested the accuracy of self-estimates, not only in psychological domains like intelligence, school achievement, creativity, or social skills but also in domains like sports or even sewing abilities. This research is well-documented in several meta-analyses (e.g., Freund and Kasten, 2012) that were ultimately integrated within a meta-synthesis (Zell and Krizan, 2014). These meta-studies found correlations between self-estimates and more "objective" measures like psychometric tests, school grades, or performance ratings from sport trainers of around only 0.3. This seems surprisingly small when compared to the often higher correlations of objective tests with external criteria like educational or professional success (e.g., around 0.5 in Schmidt and Hunter, 2004).

Self-estimates of abilities are often overly positive but sometimes also too pessimistic (see also Neubauer and Hofer, 2020). Some work investigated the sources of these individual differences. The most well-known example—the Dunning-Kruger effect (Kruger and Dunning, 1999)—suggests that especially people with low competence do not recognize their deficits. Notably, our recent findings question the generalizability of this effect (Hofer et al., 2022c; see also Gignac and Zajenkowski, 2020). Other research showed that personality traits were associated with self-estimates and their accuracy. For example, people higher in narcissism showed a higher tendency toward overestimating their abilities (e.g., Gabriel et al., 1994). Our data further indicated that self-estimates of abilities might even reflect more of a person's personality than of their "real ability" (Neubauer and Hofer, 2021; see also Herreen and Zajac, 2018).

Research seems to disagree on how detrimental inaccurate self-estimates are: Some studies found that accurate self-estimates are optimal for well-being (Kim et al., 2010; Kim and Chiu, 2011) but others reported positive (Humberg et al., 2019) or even overly positive (Dufner et al., 2018; He and Côté, 2019) estimates as more advantageous. What we and many other authors agree on is that inaccurate self-estimates could misguide important life decisions (e.g., Ackerman and Wolman, 2007; Freund and Kasten, 2012; Neubauer et al., 2018). For example, girls tend to underestimate their mathematical and visuospatial abilities, which could be one reason for why they are less likely to choose a career in a STEM field (see also Steinmayr and Spinath, 2009).

The relatively low accuracy of self-estimates begs the question what others—such as teachers, parents, or peers—know about a person's abilities. Could they help people to gain more insight into their own abilities? Indeed, the "other-perspective" is often considerably—and sometimes even surprisingly (e.g., Borkenau and Liebler, 1993)—accurate (for a review, see Neubauer and Hofer, 2020). Other-estimates have also been associated with important consequences, for example via self-fulfilling prophecies in the school context, according to which teachers' expectations of their students' intellectual potential affects students' intellectual development (Rosenthal and Jacobson, 1968; for critical review see Jussim and Harber, 2005).

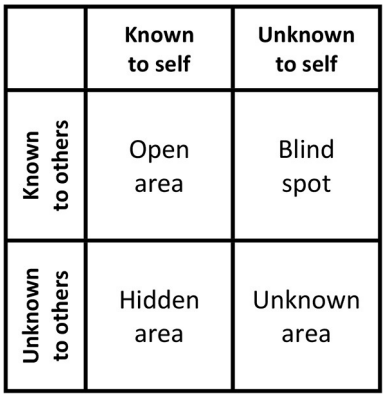


FIGURE 1
The Johari window (Luft and Ingham, 1955)/self-other knowledge asymmetry model (Vazire, 2010).

Until recently, self- and other-estimates were mostly investigated in two separate lines of research. However, the two perspectives might potentially provide different insights on ability domains and, therefore, complement each other. We compared both perspectives' accuracy in the framework of two well-known models: (1) In the Johari-window (Luft and Ingham, 1955), a trait can fall into one of four quadrants, based on whether the self, others, both perspectives, or neither perspective can assess this trait accurately (see Figure 1). (2) The self-other knowledge asymmetry (SOKA; Vazire, 2010) model is an extension of the Johari window and aims to predict personality traits' locations in the quadrants. We investigated self-other knowledge asymmetries in six central abilities: verbal, numerical, spatial intelligence, inter- and intrapersonal abilities, as well as creative potential. In a series of studies (Neubauer et al., 2018; Hofer et al., 2022a,b), we found verbal intelligence often located in the blind spot, with other persons (e.g., peers or friends) having better (i.e., more accurate) insight than the self. While particularly numerical intelligence and creativity were often in the open area (i.e., both the self and others were at least somewhat accurate), intra- and interpersonal abilities were predominantly in the hidden area (i.e., the self knew more about them than others did). Finally, in some instances, neither people themselves nor others had insight into a person's spatial intelligence, meaning that this ability was in the unknown area. Notably, we also found that what others knew about a person's abilities depended on their relationship to this person: Close others like romantic partners or friends were often more

accurate than acquaintances (e.g., work colleagues or classmates) or strangers.

Discussion

Ability estimates as a form of beliefs

In our view, self-estimates of abilities—and related constructs like self-esteem, self-concept, or self-efficacy—as well as other-estimates of abilities are conceptually close to beliefs. Seitz and Angel (2020) suggested that beliefs are characterized by four aspects:

- Humans tend to believe they are true;
- humans have a positive stance on them;
- they can be updated though new (confirming or disproving) evidence and;
- the processes behind believing are an expression of a brain function.

Thus, believing can be considered as process, a concept termed "credition" (Angel, 2013). Hans-Ferdinand Angel (2013, p. 536) states that creditions "… are connected with empathy, perception, action control, memory, and the self-concept," thus, explicitly relating creditions to the self-concept. Of course, beliefs are much more comprehensive: They can span factual, autobiographical, semantic, ethical, political, and religious domains (e.g., Seitz and Angel, 2020).

Based on different believing processes, Seitz and Angel (2020) distinguished empirical, relational, and conceptual beliefs. Empirical and relational beliefs are thought to develop instantaneously and subconsciously, whereas conceptual beliefs are considered more complex and language bound. We consider ability estimates to include aspects of all three types of beliefs. Ability estimates are empirical as they are partially inferred based on experiences. When ability estimates are made in comparison with other people (e.g., Holling and Preckel, 2005), these estimates are relational as well. Finally, when ability estimates result from abstract processing, they are similar to conceptual beliefs, which are thought to be "… ubiquitous in our cultural life and probably build the fundament for our self-understanding in our social environment …" (Seitz and Angel, 2020, p. 4).

The literature on beliefs yields further similarities to ability estimates. Just like self-estimates, beliefs are thought to guide behavior (e.g., Seitz and Angel, 2020; Seitz et al., in press). Beliefs can also be inaccurate (i.e., misbeliefs) and inaccurately positive self-estimates of abilities could be viewed as examples of misbeliefs (McKay and Dennett, 2009). Similar to inaccurate ability estimates, there also has been discourse about whether misbeliefs might be detrimental or sometimes even beneficial (see the contribution by McKay and Dennett, 2009 and its discussion in the same journal issue). Finally, beliefs are thought to be malleable: People might update them through learning (e.g., Seitz and Angel, 2020). Similarly, there is some evidence that people update their ability estimates after receiving feedback (e.g., Carpenter et al., 2019).

Future work on ability estimates and beliefs

Importantly, there are also areas where the research traditions on ability estimates and beliefs might learn from each other. As an example, the process-perspective on beliefs exemplified in the credition model does not yet seem to be well-represented in the ability estimate literature. While there is some work on the development and neural correlates of self-concepts (e.g., Chavez and Wagner, 2020; Van der Aar et al., 2022), we have yet to encounter an agreed-upon model on how people arrive at their assessments of their own and others' abilities. On the other hand, our work on ability estimates has highlighted the relevance of differentiating between ability domains and sources of estimates (i.e., the self and different types of others). Thus, future work on the intersection between ability estimates and beliefs/creditions could benefit both areas. This research could include questions from diverse fields:

- Neuroscience: Where in the brain are ability beliefs located and is this the same across ability domains (e.g., verbal vs. numerical intelligence) and sources of beliefs (e.g., self vs. other)? By using (functional) MRI, can we distinguish people who are actually gifted from those who only believe they are gifted? Conversely, can we identify "gifted underachievers," meaning individuals possessing high ability but not "believing" in it or making use of it (for earlier studies, see Staudt and Neubauer, 2006; Bergner and Neubauer, 2011).
- Genetics: As (cognitive) abilities have a strong genetic base (e.g., Plomin and von Stumm, 2018), the question arises whether believing in one's abilities might also be partially genetically driven. If so, we may ask what genes are involved in an ability *per se* vs. the belief in said ability.
- Developmental psychology: The development of (cognitive) abilities is also impacted by what people experience in their (early) lives (e.g., schooling; Ceci, 1991). Which (childhood and adolescence) experiences foster ability-related beliefs; which hinder them?
- Work and organizational psychology: What are the positive and negative effects of (overly) high ability beliefs (e.g., Humberg et al., 2019)? Could overestimating one's abilities in a certain domain bear positive achievement outcomes, e.g., by having more self-confidence, higher self-efficacy etc.?

Conclusion

How people view their own and one another's abilities could be seen as a form of beliefs. While there are many parallels between the (mostly) psychological research on ability estimates and the broader and emerging field on beliefs and creditions, there are also areas where both could learn from each other. We believe that researchers from each of these fields would benefit from knowledge of the insights gained in the other. In interdisciplinary discussions, researchers should be aware that different terminology might be applied to conceptually very similar constructs (e.g., self-estimates and other "self-variables") so that they can avoid so-called "jangle fallacies" (i.e., assuming two concepts are very different from one another when they are not; e.g., Hagger, 2014; Marsh et al., 2019). Future research on self-estimates and creditions should help to untangle similarities vs. differences of these concepts and consequently their convergent vs. discriminant validities.

Author contributions

AN conceptualized the idea behind the manuscript. AN and GH wrote sections of the manuscript. Both authors contributed to manuscript revision, read, and approved the submitted version.

Funding

This paper was funded by Rüdiger Seitz, via the Volkswagen Foundation, Siemens Healthineers, and the Betz Foundation.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Acar, S., Tadik, H., Myers, D., van der Sman, C., and Uysal, R. (2021). Creativity and well-being: a meta-analysis. *J. Creat. Behav*. 55, 738–751. doi: 10.1002/jocb.485

Ackerman, P. L., and Wolman, S. D. (2007). Determinants and validity of self-estimates of abilities and self-concept measures. *J. Exp. Psychol. Appl*. 13, 57–78. doi: 10.1037/1076-898X.13.2.57

Angel, H.-F. (2013). "Credition, the process of belief," in *Encyclopedia of Sciences and Religions*, Eds A. L. C. Runehov and L. Oviedo (Dordrecht: Springer Netherlands), 536–539.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychol. Rev*. 84, 191–215. doi: 10.1037/0033-295X.84.2.191

Bergner, S., and Neubauer, A. C. (2011). Sex and training differences in mental rotation: a behavioral and neurophysiological comparison of gifted achievers, gifted underachievers and average intelligent achievers. *High Abil. Stud*. 22, 155–177. doi: 10.1080/13598139.2011.628849

Borkenau, P., and Liebler, A. (1993). Convergence of stranger ratings of personality and intelligence with self-ratings, partner ratings, and measured intelligence. *J. Pers. Soc. Psychol*. 65, 546–553. doi: 10.1037/0022-3514.65.3.546

Carpenter, J., Sherman, M. T., Kievit, R. A., Seth, A. K., Lau, H., and Fleming, S. M. (2019). Domain-general enhancements of metacognitive ability through adaptive training. *J. Exp. Psychol. Gen*. 148, 51–64. doi: 10.1037/xge0000505

Ceci, S. J. (1991). How much does schooling influence general intelligence and its cognitive components? A reassessment of the evidence. *Dev. Psychol*. 27, 703–722. doi: 10.1037/0012-1649.27.5.703

Chavez, R. S., and Wagner, D. D. (2020). The neural representation of self is recapitulated in the brains of friends: a round-robin fMRI study. *J. Pers. Soc. Psychol*. 118, 407–416. doi: 10.1037/pspa0000178

Cogan, L., Conklin, A., and Hollingworth, H. (1915). An experimental study of self-analysis, estimates of associates, and the results of tests. *Sch. Soc*. 2, 171–179.

Cooper, C. (2020). *Individual Differences and Personality, 4th Edn*. London: Routledge.

Dufner, M., Gebauer, J. E., Sedikides, C., and Denissen, J. J. A. (2018). Self-enhancement and psychological adjustment: a meta-analytic review. *Pers. Soc. Psychol. Rev*. 23, 48–72. doi: 10.1177/1088868318756467

Freund, P. A., and Kasten, N. (2012). How smart do you think you are? A meta-analysis on the validity of self-estimates of cognitive ability. *Psychol. Bull*. 138, 296–321. doi: 10.1037/a0026556

Gabriel, M. T., Critelli, J. W., and Ee, J. S. (1994). Narcissistic illusions in self-evaluations of intelligence and attractiveness. *J. Pers*. 62, 143–155. doi: 10.1111/j.1467-6494.1994.tb00798.x

Gignac, G. E., and Zajenkowski, M. (2020). The Dunning-Kruger effect is (mostly) a statistical artefact: valid approaches to testing the hypothesis with individual differences data. *Intelligence* 80, 101449. doi: 10.1016/j.intell.2020.101449

Hagger, M. (2014). Avoiding the "déjà-variable" phenomenon: Social psychology needs more guides to constructs. *Front. Psychol*. 5:52. doi: 10.3389/fpsyg.2014.00052

Harari, M. B., Reaves, A. C., and Viswesvaran, C. (2016). Creative and innovative performance: a meta-analysis of relationships with task, citizenship, and counterproductive job performance dimensions. *Eur. J. Work Organ. Psychol*. 25, 495–511. doi: 10.1080/1359432X.2015.1134491

He, J. C., and Côté, S. (2019). Self-insight into emotional and cognitive abilities is not related to higher adjustment. *Nat. Hum. Behav*. 3, 867–884. doi: 10.1038/s41562-019-0644-0

Herreen, D., and Zajac, I. (2018). The reliability and validity of a self-report measure of cognitive abilities in older adults: more personality than cognitive function. *J. Intell*. 6, 1–15. doi: 10.3390/jintelligence6010001

Hofer, G., Langmann, L., Burkart, R., and Neubauer, A. C. (2022a). Who knows what we are good at? Unique insights of the self, knowledgeable informants, and strangers into a person's abilities. *J. Res. Pers*. 98, 104226. doi: 10.1016/j.jrp.2022.104226

Hofer, G., Macher, S., and Neubauer, A. C. (2022b). Love is not blind: what romantic partners know about our abilities compared to ourselves, our close friends, and our acquaintances. *J. Res. Pers.* 98, 104211. doi: 10.1016/j.jrp.2022.104211

Hofer, G., Mraulak, V., Grinschgl, S., and Neubauer, A. C. (2022c). Less-intelligent and unaware? Accuracy and Dunning–Kruger effects for self-estimates of different aspects of intelligence. *J. Intellig.* 10, 10. doi: 10.3390/jintelligence10010010

Holling, H., and Preckel, F. (2005). Self-estimates of intelligence—-methodological approaches and gender differences. *Pers. Individ. Dif.* 38, 503–517. doi: 10.1016/j.paid.2004.05.003

Humberg, S., Dufner, M., Schönbrodt, F. D., Geukes, K., Hutteman, R., Küfner, A. C. P., et al. (2019). Is accurate, positive, or inflated self-perception most advantageous for psychological adjustment? A competitive test of key hypotheses. *J. Pers. Soc. Psychol.* 116, 835–859. doi: 10.1037/pspp0000204

Jussim, L., and Harber, K. D. (2005). Teacher expectations and self-fulfilling prophecies: knowns and unknowns, resolved and unresolved controversies. *Pers. Soc. Psychol. Rev.* 9, 131–155. doi: 10.1207/s15327957pspr0902_3

Kim, Y.-H., Chiu, C., and Zou, Z. (2010). Know thyself: misperceptions of actual performance undermine achievement motivation, future performance, and subjective well-being. *J. Pers. Soc. Psychol.* 99, 395–409. doi: 10.1037/a0020555

Kim, Y.-H., and Chiu, C.-Y. (2011). Emotional costs of inaccurate self-assessments: both self-effacement and self-enhancement can lead to dejection. *Emotion* 11, 1096–1104. doi: 10.1037/a0025478

Kruger, J., and Dunning, D. (1999). Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *J. Pers. Soc. Psychol.* 77, 1121–1134. doi: 10.1037/0022-3514.77.6.1121

Luft, J., and Ingham, H. (1955). "The Johari window: a graphic model of interpersonal awareness," in *Proceedings of the Western Training Laboratory in Group Development* (Los Angeles, CA: UCLA).

Marsh, H. W. (1990). The structure of academic self-concept: the Marsh/Shavelson model. *J. Educ. Psychol.* 82, 623–636. doi: 10.1037/0022-0663.82.4.623

Marsh, H. W., Pekrun, R., Parker, P. D., Murayama, K., Guo, J., Dicke, T., et al. (2019). The murky distinction between self-concept and self-efficacy: beware of lurking jingle-jangle fallacies. *J. Educ. Psychol.* 111, 331–353. doi: 10.1037/edu0000281

McKay, R. T., and Dennett, D. C. (2009). The evolution of misbelief. *Behav. Brain Sci.* 32, 493–510. doi: 10.1017/S0140525X09990975

Neubauer, A. C., and Hofer, G. (2020). "Self- and other-estimates of intelligence," in *The Cambridge Handbook of Intelligence,* 2nd Edn. Ed R. J. Sternberg (Cambridge: Cambridge University Press), 1179–1200.

Neubauer, A. C., and Hofer, G. (2021). Self-estimates of abilities are a better reflection of individuals' personality traits than of their abilities and are also strong predictors of professional interests. *Pers. Individ. Dif.* 169, 109850. doi: 10.1016/j.paid.2020.109850

Neubauer, A. C., Pribil, A., Wallner, A., and Hofer, G. (2018). The self–other knowledge asymmetry in cognitive intelligence, emotional intelligence, and creativity. *Heliyon* 4, e01061. doi: 10.1016/j.heliyon.2018.e01061

Plomin, R., and von Stumm, S. (2018). The new genetics of intelligence. *Nat. Rev. Genet.* 19, 148–159. doi: 10.1038/nrg.2017.104

Rosenberg, M. (1965). *Society and the Adolescent Self-Image*. Princeton University Press. Available online at: https://search.ebscohost.com/login.aspx?direct=trueandscope=siteanddb=nlebkandAN=1078436

Rosenthal, R., and Jacobson, L. (1968). Pygmalion in the classroom. *Urban Rev.* 3, 16–20. doi: 10.1007/BF02322211

Schmidt, F. L., and Hunter, J. (2004). General mental ability in the world of work: occupational attainment and job performance. *J. Pers. Soc. Psychol.* 86, 162–173. doi: 10.1037/0022-3514.86.1.162

Seitz, R. J., and Angel, H.-F. (2020). Belief formation – a driving force for brain evolution. *Brain Cogn.* 140, 105548. doi: 10.1016/j.bandc.2020.105548

Seitz, R. J., Angel, H., and Paloutzian, R. F. (in press). Bridging the gap between believing and memory functions. *Euro. J. Psychol.* doi: 10.23668/psycharchives.5421

Staudt, B., and Neubauer, A. C. (2006). Achievement, underachievement and cortical activation: a comparative EEG study of adolescents of average and above-average intelligence. *High Abil. Stud.* 17, 3–16. doi: 10.1080/13598130600946855

Steinmayr, R., and Spinath, B. (2009). What explains boys' stronger confidence in their intelligence? *Sex Roles* 61, 736–749. doi: 10.1007/s11199-009-9675-8

Van der Aar, L. P. E., Peters, S., Becht, A. I., and Crone, E. A. (2022). Better self-concept, better future choices? Behavioral and neural changes after a naturalistic self-concept training program for adolescents. *Cogn. Affect. Behav. Neurosci.* 22, 341–361. doi: 10.3758/s13415-021-00946-1

Vazire, S. (2010). Who knows what about a person? The self–other knowledge asymmetry (SOKA) model. *J. Pers. Soc. Psychol.* 98, 281–300. doi: 10.1037/a0017908

Zell, E., and Krizan, Z. (2014). Do people have insight into their abilities? A metasynthesis. *Perspect. Psychol. Sci.* 9, 111–125. doi: 10.1177/1745691613518075

frontiers | Frontiers in Behavioral Neuroscience

TYPE Opinion
PUBLISHED 18 August 2022
DOI 10.3389/fnbeh.2022.943809

Check for updates

OPEN ACCESS

EDITED BY
Rüdiger J. Seitz,
Heinrich Heine University of Düsseldorf, Germany

REVIEWED BY
Davor Madzarevic,
University of Graz, Austria

*CORRESPONDENCE
Motoaki Sugiura
sugiura@tohoku.ac.jp

SPECIALTY SECTION
This article was submitted to Emotion Regulation and Processing, a section of the journal Frontiers in Behavioral Neuroscience

RECEIVED 14 May 2022
ACCEPTED 01 August 2022
PUBLISHED 18 August 2022

CITATION
Sugiura M (2022) Adaptability, supernaturalness, and the neurocognitive basis of the self-transcendence trait: Toward an integrated framework through disaster psychology and a self-agency model. *Front. Behav. Neurosci.* 16:943809. doi: 10.3389/fnbeh.2022.943809

COPYRIGHT
© 2022 Sugiura. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Adaptability, supernaturalness, and the neurocognitive basis of the self-transcendence trait: Toward an integrated framework through disaster psychology and a self-agency model

Motoaki Sugiura[1,2]*

[1]Institute of Development, Aging and Cancer, Tohoku University, Sendai, Japan, [2]International Research Institute of Disaster Science, Tohoku University, Sendai, Japan

KEYWORDS

adaptability, self-transcendence, spirituality, religiosity, self-agency, intentional binding, forward model, brain

Introduction

Self-transcendence (ST) refers to expansion beyond the boundaries of the self in diverse dimensions, including physical and social. It often also includes expanded, prosocial, spiritual, and religious worldviews, as well as psychological and behavioral qualities that are achieved through ST (Garcia-Romeu, 2010). For empirical research of ST as a trait, many questionnaires have been developed (Kitson et al., 2020). In particular, the Self-Transcendence Scale (STS) (Reed, 1991) proposed in the field of nursing and the ST subscale of the Temperament and Character Inventory (TCI-ST) (Cloninger et al., 1993) in the field of psychobiology have contributed significantly to this research (Garcia-Romeu, 2010). These questionnaires conceptualize ST as the final stage of human psychological development and adaptability, particularly in old age.

However, the characteristics considered by these two ST trait questionnaires are dominated either by adaptability or supernaturalness. This may hinder the integration of this line of empirical research into the theoretical literature on ST, in which the coexistence of adaptability and supernaturalness is taken for granted (Yaden et al., 2017; Kaufman, 2020). The STS was developed to measure adaptive psychological and behavioral traits in older adults in the terminal stages of illness. Its items mainly evaluate connectedness and are intuitively acceptable to most people as adaptive (Reed, 1991). Only 1 ("finding meaning in my spiritual beliefs") of 15 items has a slight supernatural nuance, which is inadequate for researchers who are interested in the relationship between the ST trait and spirituality or religiosity. Empirical studies that have used the STS have reported an association between scores and well-being in a variety of populations, including healthy young adults, and an increase in scores due to health-related vulnerability and age. Based on these findings, a model has been proposed in which ST moderates the negative impact of vulnerability on well-being (Reed, 2013). In contrast, the TCI-ST includes

many items with supernatural, spiritual, and religious nuances, probably due to the multidimensional nature of TCI and the need for differentiation from other adaptive dimensions. Because of its uniqueness, the TCI-ST has gained significant attention and has been used in many studies. There is, however, little evidence of an association between TCI-ST scores and adaptability, such as well-being (Cloninger and Zohar, 2011; Spittlehouse et al., 2014; Moreira et al., 2015). Instead, many studies have reported an association between TCI-ST scores and psychotic traits (MacDonald and Holland, 2002; Ohi et al., 2012; Gaweda et al., 2015).

Although some cognitive bias is assumed to underlie the multiple dimensions of ST, its cognitive and neural bases are unknown. Previous research has focused mainly on the supernatural aspects of ST, considering them inseparable from spirituality and religiosity (MacDonald and Holland, 2002; Urgesi et al., 2010; Kitson et al., 2020). Anthropologically, these traits are considered to be linked to a higher-level cognitive bias inherent in humans (Bulbulia, 2004; Boyer and Bergstrom, 2008), such as the imagination that enables the formation of transcendent societies based on essentialized roles and groups (Bloch, 2008). Many neuroimaging studies have addressed the neural correlates of ST in terms of the experience or trait; however, an integrated view has yet to be achieved. Studies on spiritual or religious supernatural experiences are abundant, reporting diverse and different activation areas across studies (Rim et al., 2019; Kitson et al., 2020). Two studies have addressed the trait of ST or religiosity; notably, they found associations with decreased brain activity (Kapogiannis et al., 2009) and brain damage (Urgesi et al., 2010) of partially overlapping areas.

This paper uses recent findings in disaster psychology and the neurocognitive model of self-agency to consider whether adaptability and supernaturalness coexist in ST traits, as well as to evaluate the common cognitive bias and its neural basis that underlie the multifaceted nature of the ST trait.

Do adaptability and supernaturalness coexist?

Recent disaster psychology research has identified ST trait concepts that include supernatural nuances. In a study that explored the psycho-behavioral characteristics that were advantageous for survival (Power to Live; P2L) among the survivors of the 2011 Great East Japan Earthquake (Sugiura et al., 2015a), eight factors were identified, including one that was consistent with ST, which consisted of the following four items (P2L-ST):

- I am aware that I am alive, have a sense of responsibility in living.
- I am aware of the path and teachings I should follow as a person.
- I am aware of the role I should play in society.
- I think that my actions towards others will go around and eventually come back to me.

It is noteworthy that items 1 and 4 have a supernatural nuance and overlap with the TCI-ST items. These items overlap with the representative items (i.e., with high loadings) of another ST questionnaire constructed in Japan from the viewpoint of transpersonal psychology (Nakamura, 1998).

The P2L-ST has also been demonstrated to have adaptability in terms of moderating the relationship between vulnerability and well-being, as proposed for STS in a nursing theory (Reed, 2013). The effect of vulnerability on P2L-ST seems evident, given the significant association of scores with disaster experience and age (Figure 1A) when the P2L-ST completed by members of the general population ($n = 1200$) (Ishibashi et al., 2019) was compared with data from disaster survivors ($n = 1350$) (Sugiura et al., 2015a). A three-way analysis of variance of disaster experience (two levels) × age (20s to >70s; six levels) × sex (two levels) showed medium ($\eta^2 > 0.06$) and small ($\eta^2 > 0.01$) main effect sizes for disaster experience [$F_{(1,2526)} = 190.629$, $p < 0.001$, $\eta^2 = 0.068$] and age [$F_{(5,2526)} = 10.562$, $p < 0.001$, $\eta^2 = 0.019$]; I inferenced significance using the effect size η^2 (Cohen, 1992) rather than p-value, considering the large sample size. For the further details of the data, analysis, or results, see Supplementary Tables S1, S2. A relationship between P2L-ST and well-being has also been demonstrated. P2L-ST scores are associated with housing reconstruction and well-being during the reconstruction phase in survivors who have lost their housing; this association is not observed in survivors who have not lost their housing (Sato et al., 2021). Scores have also been reported to be positively correlated with helping behavior during tsunami evacuation (Sugiura et al., 2020).

Cognitive and neural bases

Recent studies have suggested that the cognitive bias underlying the ST trait is related to the sense of self-agency based on forward model predictions. A study of the relationship between intentional binding, an established objective measure of the sense of self-agency in action, and the eight factors of the P2L identified a significant positive correlation between intentional binding and P2L-ST (Niikuni et al., 2022). Intentional binding is the process or degree of bias by which the time interval between one's action and the consequent sensory input is perceived to be shortened by forward model prediction; this is considered to be related to the sense of self-agency, particularly at the unconscious level (Haggard, 2017). Considering that intentional binding creates an, arguably illusory, consciousness of the relationship between the self and the external world, it may be a common cognitive basis of ST and a sense of self-agency. Intentional binding is also related to various adaptive traits.

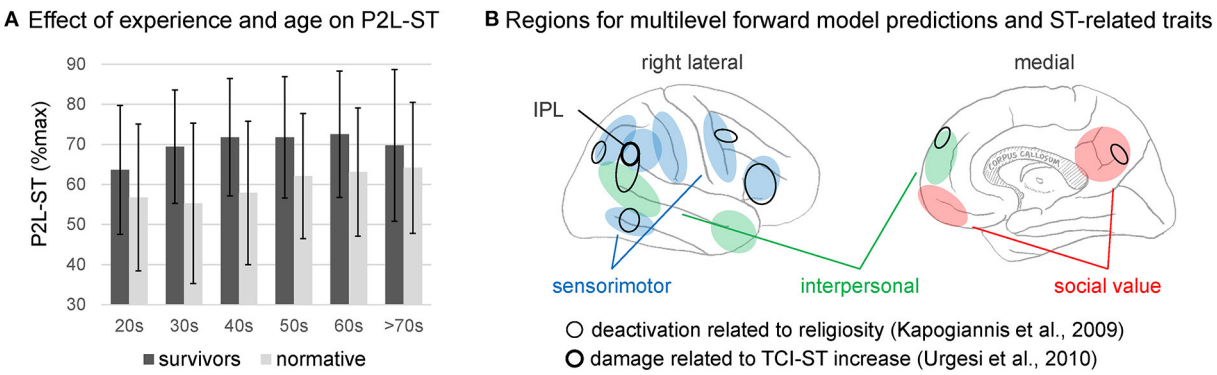


FIGURE 1
(A) Average and standard deviation (error bar) of the score (%max) of self-transcendence (ST) factor of the power to live (P2L) questionnaire separately for six age levels (from 20s to over 70) and two groups: disaster survivors (dark gray) and normative population (light gray). **(B)** Cortical regions implicated in forward model prediction in the three-layer model of self-agency (Sugiura13) and those implicated in ST-related traits. Regions are schematically drawn on the lateral (left panel) and medial (left panel) surface of the right cerebral hemisphere. Regions for the sensorimotor, interpersonal, and social value levels in the three-layer model of self-agency are given in blue, green, and red, respectively. Damage related to the increase in the ST score of the TCI [Temperament and Character Inventory Cloninger et al., 1993] (Urgesi et al., 2010) and deactivation related to a higher perceived level of God's involvement (Kapogiannis et al., 2009) are encircled by thick and thin lines, respectively.

Strong intentional binding correlates with the belief in free will (Aarts and van den Bos, 2011), while weak intentional binding is associated with various negative psychological conditions, such as schizophrenia (Graham-Schmidt et al., 2016), obsessive-compulsive tendencies (Oren et al., 2019), and narcissism (Render and Jansen, 2019). Notably, similar to the TCI-ST, an association between intentional binding and psychotic tendencies (Graham et al., 2015) has been reported.

Does the cognitive bias, which is apparently limited to the sensorimotor domain, give rise to the multidimensionality of ST, including various social domains? Recent theories of social cognition and developmental psychology allow such conceptual expansion. It has been proposed that the forward model prediction process for the sense of self-agency in action allows for the development of the ability to perceive interactional relationships between the self and others (sense of shared agency) through repeated social interactions during infancy (Gergely, 2001). In line with this, studies have demonstrated a relationship between intentional binding and the sense of shared agency (Obhi and Hall, 2011), as well as between low intentional binding and low theory of mind ability in autism spectrum disorders (Zalla et al., 2015). Furthermore, a three-layer model of the sense of self-agency (Sugiura, 2013), inspired by the theory of adolescent developmental psychology (Cooley, 1902; Mead, 1934), suggests an extension of the forward model prediction process not only from the sensorimotor (action agency) to interpersonal (shared agency) levels but also to the social-value level, which concerns the awareness of one's social role and value in the larger social context, and may be related to the prosocial and moral dimensions of ST.

This conceptual expansion appears to be supported by neurobiological findings related to the ST trait, which imply a link between the ST trait and the mechanisms that inhibit multilevel forward model prediction and resulting error detection. In two previous studies on the trait of ST or religiosity, ST was associated with reduced brain activity or brain damage. Indeed, in general, a sense of self-agency is associated with reduced activity in brain regions involved in forward model prediction or related error detection. Damage to the inferior parietal lobule (IPL) is associated with elevated TCI-ST scores (Urgesi et al., 2010); this region has been implicated in prediction or error detection at the sensorimotor level (Schnell et al., 2007; Kikuchi et al., 2019). While thinking about religious beliefs, a relatively lower degree of activation has been identified in various cortical areas in individuals with a higher perceived level of God's involvement (Kapogiannis et al., 2009). The areas distributed over the lateral and medial cortex of the right cerebral hemisphere overlap with the regions for forward model prediction at the sensorimotor level (including the IPL), as well as at higher levels (i.e., interpersonal and social values) in the three-layer model of the sense of self-agency (Sugiura, 2013) (Figure 1B).

Discussion and conclusion

The ST concept identified in recent disaster psychology research (P2L-ST) was thus found adaptive in terms of moderating the relationship between vulnerability and well-being, and included moderate supernatural nuances. The common cognitive bias underlying the multidimensionality of ST has been suggested to be related to a sense of self-agency, indicating the possibility that the bias is caused by a process

that controls the neural networks involved in multilevel forward model prediction.

The latter conceptualization may allow for the understanding of individual differences in a variety of ST-relevant beliefs, such as cultural and religious beliefs, according to a recent theoretical framework of the believing process (Sugiura et al., 2015b). The framework attributes the characteristics of the believing process (e.g., self-organization and stability) to the structure of the belief representations composed of perceptual, action, and value components; the associations between the former two make up the very basis of the forward-model prediction. This framework also assumes a hierarchically nested structure of the representations in the three levels (Sugiura, 2013). Individual conformity to supernatural beliefs may be explained by the individual strength of common cognitive bias prevalent across multi-level believing processes.

These findings and hypotheses may also facilitate anthropological discussions of the development of human-specific sociality and culture, including religion, starting from the ST trait. Supernaturalness seems to be key to relating the ST concept to unique natures of religion and culture, and adaptability is the premise for discussing it in the context of evolution and development. Future discussions are expected as to whether anthropological hypotheses on the development of human-specific societies and cultures, including religion (Bulbulia, 2004; Bloch, 2008), are consistent with the notion of common cognitive bias between ST and the sense of self-agency, and with the neurocognitive hypothesis on the notion based on multilevel forward model prediction and its control process.

Several issues remain unaddressed. First, the relationship between the ST trait and psychosis requires further investigation. Although P2L-ST has not been examined regarding this issue, intentional binding, which correlates with the ST trait, is correlated with psychosis (Graham et al., 2015). The apparently contradictory associations may be because supernatural beliefs are adaptive only at a moderate level or because supernatural beliefs are an adaptive response to internal psychological or neurological adversity. Second, relationships between diverse supernatural, mystical, and religious experiences and beliefs, which are extensively evaluated in the TCI-ST, and the adaptability identified in the P2L-ST, are also uninvestigated. The implications of the neural activity reported in various brain regions in relation to supernatural experiences and beliefs also remain to be elucidated. Finally, the process through which the ST trait is enhanced by vulnerability is unknown. The process seems to be multiphasic; in the short term, intentional binding is weakened by negative events (Obhi et al., 2013) before the facilitatory effect of vulnerability on ST becomes apparent in the long term.

Author contributions

MS conceptualized, analyzed the data, and wrote the paper.

Funding

This study was supported by KAKENHI (20K20292) from the Japan Society for the Promotion of Science, and the Subsidy for Interdisciplinary Study and Research Concerning COVID-19 from the Mitsubishi Foundation. The publication of this paper is funded by Rüdiger Seitz, *via* the Volkswagen Foundation, Siemens Healthineers, and the Betz Foundation. The funders were not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnbeh.2022.943809/full#supplementary-material

References

Aarts, H., and van den Bos, K. (2011). On the foundations of beliefs in free will: intentional binding and unconscious priming in self-agency. *Psychol. Sci.* 22, 532–537. doi: 10.1177/0956797611399294

Bloch, M. (2008). Why religion is nothing special but is central. *Philos. Trans. R. Soc. B Biol. Sci.* 363, 2055–2061. doi: 10.1098/rstb.2008.0007

Boyer, P., and Bergstrom, B. (2008). Evolutionary perspectives on religion. *Annu. Rev. Anthropol.* 37, 111–130. doi: 10.1146/annurev.anthro.37.081407.085201

Bulbulia, J. (2004). The cognitive and evolutionary psychology of religion. *Biol. Philos.* 19, 655–686. doi: 10.1007/s10539-005-5568-6

Cloninger, C. R., Svrakic, D. M., and Przybeck, T. R. (1993). A psychobiological model of temperament and character. *Arch. Gen. Psychiatry* 50, 975–990. doi: 10.1001/archpsyc.1993.01820240059008

Cloninger, C. R., and Zohar, A. H. (2011). Personality and the perception of health and happiness. *J. Affect. Disord.* 128, 24–32. doi: 10.1016/j.jad.2010.06.012

Cohen, J. (1992). Quantitative methods in psychology: a power primer. *Psychol. Bull.* 112, 155–159. doi: 10.1037/0033-2909.112.1.155

Cooley, C. H. (1902). *Human Nature and the Social Order*. New York, NY: Scribner.

Garcia-Romeu, A. (2010). Self-transcendence as a measurable transpersonal construct. *J. Transpers. Psychol.* 42, 26–47. Available online at: https://www.transpersonaljournal.com/en/

Gaweda, Ł., Prochwicz, K., and Cella, M. (2015). Cognitive biases mediate the relationship between temperament and character and psychotic-like experiences in healthy adults. *Psychiatry Res.* 225, 50–57. doi: 10.1016/j.psychres.2014.10.006

Gergely, G. (2001). The obscure object of desire- *'Nearly, but clearly not, like me'*: contingency preference in normal children versus children with autism. *Bull. Menninger Clin.* 65, 411–426. doi: 10.1521/bumc.65.3.411.19853

Graham, K. T., Martin-Iverson, M. T., and Waters, F. A. V. (2015). Intentional binding or perceptual repulsion? Binding in a general population sample decreases with age and increases with psychosis-like experiences. *Psychol. Conscious.* 2, 269–282. doi: 10.1037/cns0000067

Graham-Schmidt, K. T., Martin-Iverson, M. T., Holmes, N. P., and Waters, F. A. V. (2016). When one's sense of agency goes wrong: absent modulation of time perception by voluntary actions and reduction of perceived length of intervals in passivity symptoms in schizophrenia. *Conscious. Cogn.* 45, 9–23. doi: 10.1016/j.concog.2016.08.006

Haggard, P. (2017). Sense of agency in the human brain. *Nat. Rev. Neurosci.* 18, 196–207. doi: 10.1038/nrn.2017.14

Ishibashi, R., Nouchi, R., Honda, A., Abe, T., and Sugiura, M. (2019). A concise psychometric tool to measure personal characteristics for surviving natural disasters: development of a 16-item power to live questionnaire. *Geosciences* 9, 366. doi: 10.3390/geosciences9090366

Kapogiannis, D., Barbey, A. K., Su, M., Zamboni, G., Krueger, F., and Grafman, J. (2009). Cognitive and neural foundations of religious belief. *Proc. Natl. Acad. Sci. U.S.A.* 106, 4876–4881. doi: 10.1073/pnas.0811717106

Kaufman, S. B. (2020). *Transcend: The New Science of Self-Actualization*. New York, NY: Tarcher Perigee.

Kikuchi, T., Sugiura, M., Yamamoto, Y., Sasaki, Y., Hanawa, S., Sakuma, A., et al. (2019). Neural responses to action contingency error in different cortical areas are attributable to forward prediction or sensory processing. *Sci. Rep.* 9:9847. doi: 10.1038/s41598-019-46350-1

Kitson, A., Chirico, A., Gaggioli, A., and Riecke, B. E. (2020). A review on research and evaluation methods for investigating self-transcendence. *Front. Psychol.* 11:547687. doi: 10.3389/fpsyg.2020.547687

MacDonald, D. A., and Holland, D. (2002). Examination of the psychometric properties of the temperament and character inventory self-transcendence dimension. *Pers. Individ. Dif.* 32, 1013–1027. doi: 10.1016/S0191-8869(01)00107-6

Mead, G. H. (1934). *Mind, Self, and Society from the Standpoint of a Social Behaviorist*. Chicago, IL: University of Chicago Press.

Moreira, P. A. S., Cloninger, C. R., Dinis, L., Sá, L., Oliveira, J. T., Dias, A., et al. (2015). Personality and well-being in adolescents. *Front. Psychol.* 5:1494. doi: 10.3389/fpsyg.2014.01494

Nakamura, M. (1998). A study on self-transcendence and subjective well-being: a trial of constructing the tendency for self-transcendence scale. *Bull. Fac. Educ. Ehime Univ.* 45, 59–79.

Niikuni, K., Nakanishi, M., and Sugiura, M. (2022). Intentional binding and self-transcendence: searching for pro-survival behavior in sense-of-agency. *Conscious. Cogn.* 102:103351. doi: 10.1016/j.concog.2022.103351

Obhi, S. S., and Hall, P. (2011). Sense of agency and intentional binding in joint action. *Exp. Brain Res.* 211, 655–662. doi: 10.1007/s00221-011-2675-2

Obhi, S. S., Swiderski, K. M., and Farquhar, R. (2013). Activating memories of depression alters the experience of voluntary action. *Exp. Brain Res.* 229, 497–506. doi: 10.1007/s00221-012-3372-5

Ohi, K., Hashimoto, R., Yasuda, Y., Fukumoto, M., Yamamori, H., Iwase, M., et al. (2012). Personality traits and schizophrenia: evidence from a case–control study and meta-analysis. *Psychiatry Res.* 198, 7–11. doi: 10.1016/j.psychres.2011.12.018

Oren, E., Eitam, B., and Dar, R. (2019). Intentional binding and obsessive-compulsive tendencies: a dissociation between indirect and direct measures of the sense of agency. *J. Obsessive Compuls. Relat. Disord.* 20, 59–65. doi: 10.1016/j.jocrd.2017.11.002

Reed, P. G. (1991). Self-transcendence and mental health in oldest-old adults. *Nurs. Res.* 40, 5–11. doi: 10.1097/00006199-199101000-00002

Reed, P. G. (2013). "Theory of self-transcendence," in *Middle Range Theory for Nursing, 3rd Edn* (New York, NY: Springer publishing company), 109–139. doi: 10.1891/9780826195524.0006

Render, A., and Jansen, P. (2019). Dopamine and sense of agency: determinants in personality and substance use. *PLoS ONE* 14:e0214069. doi: 10.1371/journal.pone.0214069

Rim, J. I., Ojeda, J. C., Svob, C., Kayser, J., Drews, E., Kim, Y., et al. (2019). Current understanding of religion, spirituality, and their neurobiological correlates. *Harv. Rev. Psychiatry* 27, 303–316. doi: 10.1097/HRP.0000000000000232

Sato, S., Ishibashi, R., and Sugiura, M. (2021). Two major elements of life recovery after a disaster: their impacts dependent on housing damage and the contributions of psycho-behavioral factors. *J. Disaster Res.* 16, 1107–1120. doi: 10.20965/jdr.2021.p1107

Schnell, K., Heekeren, K., Schnitker, R., Daumann, J., Weber, J., Heßelmann, V., et al. (2007). An fMRI approach to particularize the frontoparietal network for visuomotor action monitoring: detection of incongruence between test subjects' actions and resulting perceptions. *Neuroimage* 34, 332–341. doi: 10.1016/j.neuroimage.2006.08.027

Spittlehouse, J. K., Vierck, E., Pearson, J. F., and Joyce, P. R. (2014). Temperament and character as determinants of well-being. *Compr. Psychiatry* 55, 1679–1687. doi: 10.1016/j.comppsych.2014.06.011

Sugiura, M. (2013). Associative account of self-cognition: extended forward model and multi-layer structure. *Front. Hum. Neurosci.* 7:535. doi: 10.3389/fnhum.2013.00535

Sugiura, M., Nouchi, R., Honda, A., Sato, S., Abe, T., and Imamura, F. (2020). Survival-oriented personality factors are associated with various types of social support in an emergency disaster situation. *PLoS ONE* 15:e0228875. doi: 10.1371/journal.pone.0228875

Sugiura, M., Sato, S., Nouchi, R., Honda, A., Abe, T., Muramoto, T., et al. (2015a). Eight personal characteristics associated with the power to live with disasters as indicated by survivors of the 2011 Great East Japan earthquake disaster. *PLoS ONE* 10:e0130349. doi: 10.1371/journal.pone.0130349

Sugiura, M., Seitz, R. J., and Angel, H.-F. (2015b). Models and neural bases of the believing process. *J. Behav. Brain Sci.* 05, 12–23. doi: 10.4236/jbbs.2015.51002

Urgesi, C., Aglioti, S. M., Skrap, M., and Fabbro, F. (2010). The spiritual brain: selective cortical lesions modulate human self-transcendence. *Neuron* 65, 309–319. doi: 10.1016/j.neuron.2010.01.026

Yaden, D. B., Haidt, J., Hood, R. W., Vago, D. R., and Newberg, A. B. (2017). The varieties of self-transcendent experience. *Rev. Gen. Psychol.* 21, 143–160. doi: 10.1037/gpr0000102

Zalla, T., Miele, D., Leboyer, M., and Metcalfe, J. (2015). Metacognition of agency and theory of mind in adults with high functioning autism. *Conscious. Cogn.* 31, 126–138. doi: 10.1016/j.concog.2014.11.001

frontiers | Frontiers in Psychology

TYPE Opinion
PUBLISHED 13 October 2022
DOI 10.3389/fpsyg.2022.983592

Check for updates

OPEN ACCESS

EDITED BY
Rüdiger J. Seitz,
Heinrich Heine University of Düsseldorf, Germany

REVIEWED BY
Julia U. Henschke,
University Hospital Magdeburg, Germany
Patrizia Fattori,
University of Bologna, Italy

*CORRESPONDENCE
Rafael Bretas
rafael.bretasvieira@riken.jp
Atsushi Iriki
atsushi.iriki@riken.jp

SPECIALTY SECTION
This article was submitted to Theoretical and Philosophical Psychology, a section of the journal Frontiers in Psychology

RECEIVED 04 July 2022
ACCEPTED 30 September 2022
PUBLISHED 13 October 2022

CITATION
Bretas R, Tia B, Yamazaki Y and Iriki A (2022) Multisensory integration and belief in the self. *Front. Psychol.* 13:983592. doi: 10.3389/fpsyg.2022.983592

COPYRIGHT
© 2022 Bretas, Tia, Yamazaki and Iriki. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Multisensory integration and belief in the self

Rafael Bretas*, Banty Tia, Yumiko Yamazaki and Atsushi Iriki*

Laboratory for Symbolic Cognitive Development, RIKEN Center for Biosystems Dynamics Research, Kobe, Japan

KEYWORDS

somatosensory integration, mirror self recognition, consciousness, credition, beliefs, parietal cortex, self recognition

Introduction

Our first experience of the world originates from the information we receive through the senses, allowing us to make mental representations of the features that can be experienced from each part of the environment—be those objects, events, places, or beings. However, these parts are not perceived separately through each sense. Rather, sight, touch, smell, hearing, and taste are integrated early in life in multimodal areas in the brain (Lewis and Essen, 2000). While this process, together with memory, supports the formation of beliefs of increasing complexity, it is also constantly modified by those same beliefs. In this opinion paper we briefly describe some of the neural underpinnings of conscious perception and illustrate how a complex belief is formed from sensory information using the example of mirror self-recognition in macaques.

From when a sense is raised to awareness until when it is integrated into other senses, a separate process occurs. A relational association is established, one in which the codependency of these stimuli becomes their own defining characteristic. That is, an object or event is recognized by simultaneously eliciting different modalities of sensation (Crick and Koch, 1990; Deroy et al., 2016). It is important to clarify that multisensory integration does not necessarily induce a conscious process. However, unconscious integration seems to be possible in only limited conditions, such as in simple forms of visual adaptation or when a stimuli pair has been previously learned (Faivre et al., 2014; Mudrik et al., 2014).

In this sense, integration can be understood as an antecedent to a behavior, perhaps similarly to how attitudes or mindsets are modulated (Seitz and Angel, 2012) or as an *empirical belief* (Seitz and Angel, 2020). Nevertheless, becoming aware of a percept as an amalgam of sensations forms the basis for a conscious belief that can be expressed as a decision or action, or in declarative form as a *conceptual belief* (Seitz and Angel, 2020). From the point in which a sense is raised to awareness, recognition may take place. Although often understood as a single behavioral phenomenon, "recognition" arises from separate neurophysiological processes that can function independently of each other. Here, we will focus on two of these general processes that are often taken as determinators of recognition.

First, there is a memory component that locates the sensory information in the place and context where sensation occurs (Mandler, 1980), supported by connections between sensory areas in the neocortex, perirhinal cortex, and parahippocampal regions (Brown and Aggleton, 2001; Eichenbaum et al., 2007). This system, or systems, as it could be subdivided in two main components, is responsible for retrieval of contextual information and the recollection of the stimulus (Brown and Aggleton, 2001). Autobiographical memory also centers sensory experiences around oneself as the individual agent of sensation, perhaps mediated by connections between the posterior cingulate and medial parietal cortex (Rolls, 2022). The hippocampal and parahippocampal regions show extensive connections to sensory and motor areas, but despite playing a fundamental role in recollecting and situating the sensory information received in time and space, recognition itself seems to be formed independently from these regions. Moreover, lesion experiments confirm that memory is not necessary for simple object recognition (DeCoteau and Kesner, 1998; Burwell, 2000; Save and Poucet, 2000; Langston and Wood, 2010), instead it may represent the emotional contents and semantic information rather than the physical properties that allow the conscious perception of an object (Rolls, 2022; Rolls et al., 2022).

The second function that supports recognition is multisensory integration. The perception of simple physical features, such as shape, color, or texture, can be accomplished by unimodal tactile and visual processing streams without reaching awareness. Unconscious perception is also common in multimodal areas for the purposes of guiding motor control (Milner and Goodale, 2008; Mudrik et al., 2014). However, recognition (i.e., the conscious perceptual experience that allows the identification of an object or scene) recruits large, distributed networks that integrate different senses (Dijkerman and de Haan, 2007; Winters and Reid, 2010). The parietal cortex appears to be the source of this conscious perception process, being well interconnected with prefrontal, cingulate and primary sensory areas (Lewis and Essen, 2000; Vincent et al., 2006; Whitlock et al., 2008; Rolls et al., 2022). Furthermore, the parietal cortex, together with the prefrontal cortex, directs attention and modulates perception and the emotional significance of sensory events (Mesulam, 1998; Steinmetz et al., 2000; Culham and Kanwisher, 2001; Galletti et al., 2010). In humans, the parietal cortex is also functionally interconnected with language and declarative memory areas, a hallmark of conscious perception (Rolls et al., 2022).

This wide network that combines perception and memory to contextualize what is perceived may form a part of the broader consciousness, allowing one to recognize the world and the self (Dehaene and Changeux, 2011; Mudrik et al., 2014). The large-scale synchronization between sensory, motor and executive functions required for conscious perception could be mediated by the claustrum, as an area with reciprocal connections to most regions of the cortex, perhaps as a conductor of sensory experiences (Crick and Koch, 2005), with this combination process being permeated by different degrees of beliefs (Seitz and Angel, 2012).

Here, we would like to focus on the formation and modulation of some simple beliefs related to the self and the world. For example, when there is a mismatch between the senses, such as in the ventriloquist effect, localizing the source of a sound together with a movement (mouth movement, in this example) uses a combination of both auditory and visual senses, but viewers' sensations are distorted by beliefs of how reality *should* be according to previous experiences (Alais and Burr, 2004; Seitz and Angel, 2020). The viewer knows that the sound comes from the ventriloquist, not the puppet, but at the same time they also have experienced a reality in which sounds usually come from moving parts, and in particular, voices come from moving mouths.

This type of sensory conflict is clearer in the rubber hand illusion. Even when there is a stable, conscious belief that the hand seen is not one's own, a false belief associated with a recent sensation may override the belief. The effect is strong enough to cause the activation of somatosensory areas in the brain when the false hand is stroked (Ehrsson et al., 2005). Such cases could be an example of how a conscious declarative belief can be distorted by a sensory belief. In both cases the causes could maybe be reduced to the single, deeply rooted, unconscious belief intrinsic to the senses involved that sight has higher spatial accuracy than sound or touch. Therefore, sight *should* be more reliable when incongruent spatial judgements are involved.

There are also cases where sensation is not inconsistent but instead ambiguous, such as in multistable perception (i.e., when two concurrent percepts spontaneously change). In these situations, prior beliefs may act in harmony with newly acquired sensory information to guide attention and modulate perception (Sterzer et al., 2009; Conrad et al., 2012), with the speed and often mutually exclusiveness of these changes being noteworthy.

The self in the mirror

Self-recognition in the mirror was proposed by Gallup in 1977 as "a technique for providing empirical and operational substance to the existence of self-awareness" (Gallup, 1977). Despite debates over how much self-awareness intersects with mirror self-recognition, the mirror offers the perfect example of the development of different categories of belief and how they interfere with each other according to the criteria of Seitz and Angel (2020). Since humans appear to develop or be guided into mirror self-recognition very early in life, it may be difficult to imagine its emergence. Macaques however, although not having innate self-recognition, can acquire it through training and habituation (Chang et al., 2017; Bretas et al., 2021). Therefore, below we follow the trajectory of a

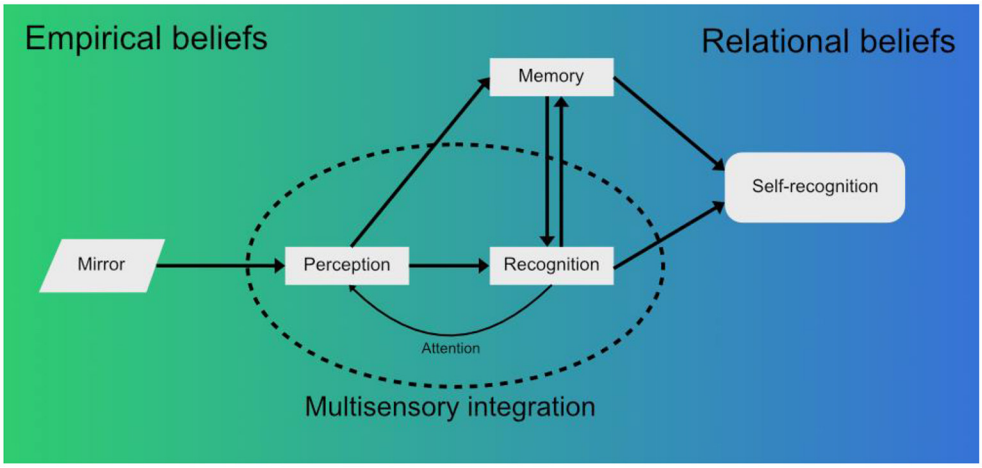


FIGURE 1
When a macaque looks into a mirror, perception of the raw visual information from the mirror forms a first approximation of reality in the form of *empirical beliefs* (Seitz and Angel, 2020). This visual stimulus is integrated with other sensory modalities, such as the proprioception of the macaque's own limbs moving, and raised to conscious awareness. Recognition of the mirror image then takes place - initially as another macaque, not the self. Simultaneously, the novelty, emotional significance and other aspects associated with the stimuli are recollected from memory—previous experiences with other macaques, the novelty of the individual seen in the mirror, innate fear, etc. The associations between what is perceived and the environment form *relational beliefs* (Seitz and Angel, 2020). These beliefs are being updated constantly according to new sensory information received from the environment, which also feedbacks into perception through attention mechanisms. Finally, a complex belief that matches the perceived stimuli may evolve, the belief that the macaque in the mirror is a reflection of the self. This belief is expressed in the form of mirror self-recognition behaviors (Chang et al., 2017; Bretas et al., 2021), but in humans it could further develop into *conceptual beliefs*, discrete, language bound concepts (Seitz and Angel, 2020) (e.g.: "I am the person in the mirror and I appear to others are they appear to me").

hypothetical macaque acquiring such a skill based on our own experimental observations (Bretas et al., 2021).

At the lowest level, there is the visual stimulus imparted by the mirror as the sensory reality (i.e., raw information that contains the visual features of the subject in the mirror). Innate visual mechanisms combined with the memory mechanisms described before allow for the classification of this stimulus (*The image in the mirror moves. It's a live being. It is a monkey.*) and its valuation in terms of emotional loading (*Is it a threat? Is it a partner?*). This process can start before the stimulus is raised to conscious awareness and is an example of *empirical beliefs* (*Figure 1*).

Before self-awareness, one must develop other-awareness, since both processes require the capacity of secondary representation (Asendorpf and Baudonnière, 1993). Both other beings and the self-body are directly accessible through the senses, but even accounting for physical similarities, there is a significant difference in perspective. At the most basic level, the spatial organization of the body will appear to be different, with one not being able to see their own face or back, for example. There is also a difference in the integration of the senses: visually the self and the other can be equally perceived, but touch is limited to the self. For example, while two people can see each other touching an object, the touch sensation can only be felt by the agent. In the same way that self-body awareness arises from the association between different unimodal sensations, such as vision and touch (Apps and Tsakiris, 2014), mirror-self recognition relies on these same sensory mechanisms as precursors.

When the macaque looks into the mirror, a third-person view of the self is shown. This view elicits new beliefs, both conflicting and ambiguous, as described in the previous section. While the body in the mirror appears from the same perspective as that of others, its movements can be precisely controlled, which is a sense of agency over a distant subject. The same could be said about the touch sensation, which now can be felt by both the agent and the mirror-image. The subject now may recollect that the primate in the mirror is always the same subject with the same face. It also does not feel like the primate's previous experiences or what is expected of the related visual stimuli. The mirror shows soft fur, but it is cold and hard to the touch; the primate in the mirror cannot be touched. These new sensorial experiences accumulate and generate *relational beliefs* about the relations with the environment. According to Sugiura et al. (2015), mental representations are formed through the association between an action and its consequential perception learned through repeated experience.

Multimodal visuo-somatosensory neurons are often spatially tunned to represent the space around the subject from an egocentric point of view, mapping the position of the own body and reachable objects nearby. These neurons estimate and guide limb movement as well as tracking objects

moving toward or nearby the subject (Taoka et al., 2013; Hihara et al., 2015; Galletti et al., 2022). But when movement is performed in front of the mirror, the mismatch between the prediction and the actual sensation may no longer appropriately represent the actual bodily or environmental state (Sugiura et al., 2015). Perception may then be updated by this new incoming sensory data and combined with past outputs and decisions to account for this new state, in accordance with the free-energy principle (Apps and Tsakiris, 2014); *relational beliefs* can alter *empirical beliefs*. This update to sensation may lead to the mirror being ignored from that point on as a useless social cue, with no new attempts to interact with this primate in the mirror and no emotional reactions of fear or dominance upon seeing it. However, discrepancy from previous beliefs in this new information received from unimodal areas could be explained away by multisensory integration, giving rise to self-recognition (Apps and Tsakiris, 2014; Chang et al., 2017; Bretas et al., 2021). Indeed, binding different sensory aspects of and object in a mutually coherent way provides the experience of perceptual unity necessary to group the individual body parts in a concept of an indivisible self-body (Crick and Koch, 1990; Bretas et al., 2021).

The acknowledgment of the self in third-person promotes the belief that the other is like the self, with empathy and the emotional valence of the new beliefs further shaping both mirror-perception and own-perception (Gallup, 1998; Sugiura et al., 2015; Bretas et al., 2020). Primates are special in that their brains evolved with the addition of new functional subdivisions to the neocortex (Dooley and Krubitzer, 2013). Areas in the parietal cortex related to self-awareness and social-awareness may be essential to the development of mental models of introspectively based social strategies and language, forming the basis for *conceptual beliefs* and culture (Gallup, 1998; Sugiura et al., 2015; Bretas et al., 2020, 2021; Seitz and Angel, 2020; Seitz, 2022). *Conceptual beliefs*, thereupon, give support to meta-beliefs, elevating relational beliefs about the other to language-bound discrete concepts (e.g., "the other believes…") to achieve a comprehensive notion of the world grounded in internal representations of the physical, social, and cultural environments (Angel and Seitz, 2016; Bretas et al., 2020; Seitz and Angel, 2020).

Author contributions

RB and AI contributed to conception and theory of the study. RB wrote the first draft of the manuscript. BT, YY, and AI wrote sections of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

Funding

The authors declare that this study received funding from Dr. Rüdiger Seitz, via the Volkswagen Foundation, Siemens Healthineers, and the Betz Foundation. The funders were not involved in the study design, collection, analysis, interpretation of data, the writing of this article, or the decision to submit it for publication.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Alais, D., and Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Curr. Biol.* 14, 257–262. doi: 10.1016/j.cub.2004.01.029

Angel, H.-F., and Seitz, R. J. (2016). Process of believing as fundamental brain function: the concept of credition. *SFU Res. Bull. Nr.* 1, 1–20. doi: 10.15135/2016.4.1.1-20

Apps, M. A. J., and Tsakiris, M. (2014). The free-energy self: a predictive coding account of self-recognition. *Neurosci. Biobehav. Rev.* 41, 85–97. doi: 10.1016/j.neubiorev.2013.01.029

Asendorpf, J. B., and Baudonnière, P.-M. (1993). Self-awareness and other-awareness: mirror self-recognition and synchronic imitation among unfamiliar peers. *Dev. Psychol.* 29, 88–95. doi: 10.1037/0012-1649.29.1.88

Bretas, R., Taoka, M., Hihara, S., Cleeremans, A., and Iriki, A. (2021). Neural evidence of mirror self-recognition in the secondary somatosensory cortex of macaque: observations from a single-cell recording experiment and implications for consciousness. *Brain Sci.* 11, 157. doi: 10.3390/brainsci11020157

Bretas, R., Taoka, M., Suzuki, H., and Iriki, A. (2020). Secondary somatosensory cortex of primates: Beyond body maps, toward conscious self-in-the-world maps. *Exp. Brain Res.* 238, 259–272. doi: 10.1007/s00221-020-05727-9

Brown, M. W., and Aggleton, J. P. (2001). Recognition memory: what are the roles of the perirhinal cortex and hippocampus? *Nat. Rev. Neurosci.* 2, 51–61. doi: 10.1038/35049064

Burwell, R. D. (2000). The parahippocampal region: corticocortical connectivity. *Ann. N. Y. Acad. Sci.* 911, 25–42. doi: 10.1111/j.1749-6632.2000.tb06717.x

Chang, L., Zhang, S., Poo, M., and Gong, N. (2017). Spontaneous expression of mirror self-recognition in monkeys after learning precise visual-proprioceptive association for mirror images. *Proc. Natl. Acad. Sci.* 114, 3258–3263. doi: 10.1073/pnas.1620764114

Conrad, V., Vitello, M. P., and Noppeney, U. (2012). Interactions between apparent motion rivalry in vision and touch. *Psychol. Sci.* 23, 940–948. doi: 10.1177/0956797612438735

Crick, F., and Koch, C. (1990). Towards a neurobiological theory of consciousness. *Semin. Neurosci.* 2, 263–275.

Crick, F. C., and Koch, C. (2005). What is the function of the claustrum? *Philos. Trans. R. Soc. B Biol. Sci.* 360, 1271–1279. doi: 10.1098/rstb.2005.1661

Culham, J. C., and Kanwisher, N. G. (2001). Neuroimaging of cognitive functions in human parietal cortex. *Curr. Opin. Neurobiol.* 11, 157–163. doi: 10.1016/S0959-4388(00)00191-4

DeCoteau, W., and Kesner, R. (1998). Effects of hippocampal and parietal cortex lesions on the processing of multiple-object scenes. *Behav. Neurosci.* 112, 68–82. doi: 10.1037/0735-7044.112.1.68

Dehaene, S., and Changeux, J.-P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron* 70, 200–227. doi: 10.1016/j.neuron.2011.03.018

Deroy, O., Faivre, N., Lunghi, C., Spence, C., Aller, M., and Noppeney, U. (2016). The complex interplay between multisensory integration and perceptual awareness. *Multisen. Res.* 29, 585–606. doi: 10.1163/22134808-00002529

Dijkerman, H. C., and de Haan, E. H. F. (2007). Somatosensory processes subserving perception and action. *Behav. Brain Sci.* 30, 189–201. doi: 10.1017/S0140525X07001392

Dooley, J., and Krubitzer, L. (2013). Cortical plasticity within and across lifetimes: How can development inform us about phenotypic transformations? *Front. Hum. Neurosci.* 7, 620. doi: 10.3389/fnhum.2013.00620

Ehrsson, H. H., Holmes, N. P., and Passingham, R. E. (2005). Touching a rubber hand: Feeling of body ownership is associated with activity in multisensory brain areas. *J. Neurosci.* 25, 10564–10573. doi: 10.1523/JNEUROSCI.0800-05.2005

Eichenbaum, H., Yonelinas, A. R., and Ranganath, C. (2007). The Medial Temporal Lobe and Recognition Memory. *Ann. Rev. Neurosci.* 30, 123–152. doi: 10.1146/annurev.neuro.30.051606.094328

Faivre, N., Mudrik, L., Schwartz, N., and Koch, C. (2014). Multisensory integration in complete unawareness: evidence from audiovisual congruency priming. *Psychol. Sci.* 25, 2006–2016. doi: 10.1177/0956797614547916

Galletti, C., Breveglieri, R., Lappe, M., Bosco, A., Ciavarro, M., and Fattori, P. (2010). Covert shift of attention modulates the ongoing neural activity in a reaching area of the macaque dorsomedial visual stream. *PLoS ONE* 5, e15078. doi: 10.1371/journal.pone.0015078

Galletti, C., Gamberini, M., and Fattori, P. (2022). The posterior parietal area V6A: An attentionally-modulated visuomotor region involved in the control of reach-to-grasp action. *Neurosci. Biobehav. Rev.* 141, 104823. doi: 10.1016/j.neubiorev.2022.104823

Gallup, G. G. (1977). Self recognition in primates: A comparative approach to the bidirectional properties of consciousness. *Am. Psychol.* 32, 329–338. doi: 10.1037/0003-066X.32.5.329

Gallup, G. G. (1998). Self-awareness and the evolution of social intelligence. *Behav. Processes* 42, 239–247. doi: 10.1016/S0376-6357(97)00079-X

Hihara, S., Taoka, M., Tanaka, M., and Iriki, A. (2015). Visual responsiveness of neurons in the secondary somatosensory area and its surrounding parietal operculum regions in awake macaque monkeys. *Cereb. Cortex* 25, 4535–4550. doi: 10.1093/cercor/bhv095

Langston, R. F., and Wood, E. R. (2010). Associative recognition and the hippocampus: Differential effects of hippocampal lesions on object-place, object-context and object-place-context memory. *Hippocampus* 20, 1139–1153. doi: 10.1002/hipo.20714

Lewis, J. W., and Essen, D. C. V. (2000). Corticocortical connections of visual, sensorimotor, and multimodal processing areas in the parietal lobe of the macaque monkey. *J. Comp. Neurol.* 428, 112–137. doi: 10.1002/1096-9861(20001204)428:1<112::aid-cne8>3.0.co;2-9

Mandler, G. (1980). Recognizing: the judgment of previous occurrence. *Psychol. Rev.* 87, 252–271. doi: 10.1037/0033-295X.87.3.252

Mesulam, M. (1998). From sensation to cognition. *Brain* 121, 1013–1052. doi: 10.1093/brain/121.6.1013

Milner, A. D., and Goodale, M. A. (2008). Two visual systems re-viewed. *Neuropsychologia* 46, 774–785. doi: 10.1016/j.neuropsychologia.2007.10.005

Mudrik, L., Faivre, N., and Koch, C. (2014). Information integration without awareness. *Trends Cogn. Sci.* 18, 488–496. doi: 10.1016/j.tics.2014.04.009

Rolls, E. T. (2022). The hippocampus, ventromedial prefrontal cortex, and episodic and semantic memory. *Prog. Neurobiol.* 217, 102334. doi: 10.1016/j.pneurobio.2022.102334

Rolls, E. T., Deco, G., Huang, C.-C., and Feng, J. (2022). The human posterior parietal cortex: Effective connectome, and its relation to function. *Cereb. Cortex* bhac266. doi: 10.1093/cercor/bhac266

Save, E., and Poucet, B. (2000). Involvement of the hippocampus and associative parietal cortex in the use of proximal and distal landmarks for navigation. *Behav. Brain Res.* 109, 195–206. doi: 10.1016/S0166-4328(99)00173-4

Seitz, R. J. (2022). Believing and beliefs—neurophysiological underpinnings. *Front. Behav. Neurosci.* 16, 880504. doi: 10.3389/fnbeh.2022.880504

Seitz, R. J., and Angel, H.-F. (2012). Processes of believing—a review and conceptual account. *Rev. Neurosci.* 23, 303–309. doi: 10.1515/revneuro-2012-0034

Seitz, R. J., and Angel, H.-F. (2020). Belief formation – a driving force for brain evolution. *Brain Cogn.* 140, 105548. doi: 10.1016/j.bandc.2020.105548

Steinmetz, P. N., Roy, A., Fitzgerald, P. J., Hsiao, S. S., Johnson, K. O., and Niebur, E. (2000). Attention modulates synchronized neuronal firing in primate somatosensory cortex. *Nature* 404, 187. doi: 10.1038/35004588

Sterzer, P., Kleinschmidt, A., and Rees, G. (2009). The neural bases of multistable perception. *Trends Cogn. Sci.* 13, 310–318. doi: 10.1016/j.tics.2009.04.006

Sugiura, M., Seitz, R. J., and Angel, H.-F. (2015). Models and neural bases of the believing process. *J. Behav. Brain Sci.* 5, 12. doi: 10.4236/jbbs.2015.51002

Taoka, M., Tanaka, M., Hihara, S., Ojima, H., and Iriki, A. (2013). Neural response to movement of the hand and mouth in the secondary somatosensory cortex of Japanese monkeys during a simple feeding task. *Somatosens. Motor Res.* 30, 140–152. doi: 10.3109/08990220.2013.779246

Vincent, J. L., Snyder, A. Z., Fox, M. D., Shannon, B. J., Andrews, J. R., Raichle, M. E., et al. (2006). Coherent spontaneous activity identifies a hippocampal-parietal memory network. *J. Neurophysiol.* 96, 3517–3531. doi: 10.1152/jn.00048.2006

Whitlock, J. R., Sutherland, R. J., Witter, M. P., Moser, M.-B., and Moser, E. I. (2008). Navigating from hippocampus to parietal cortex. *Proc. Natl. Acad. Sci.* 105, 14755–14762. doi: 10.1073/pnas.0804216105

Winters, B. D., and Reid, J. M. (2010). A distributed cortical representation underlies crossmodal object recognition in rats. *J. Neurosci.* 30, 6253–6261. doi: 10.1523/JNEUROSCI.6073-09.2010

frontiers | Frontiers in Psychology

TYPE Opinion
PUBLISHED 10 November 2022
DOI 10.3389/fpsyg.2022.951248

Check for updates

OPEN ACCESS

EDITED BY
Rüdiger J. Seitz,
Heinrich Heine University of Düsseldorf, Germany

REVIEWED BY
Giovanni Buccino,
Ospedale San Raffaele (IRCCS), Italy

*CORRESPONDENCE
Simone G. Shamay-Tsoory
sshamay@psy.haifa.ac.il

SPECIALTY SECTION
This article was submitted to Cognition, a section of the journal Frontiers in Psychology

RECEIVED 23 May 2022
ACCEPTED 01 September 2022
PUBLISHED 10 November 2022

CITATION
Shamay-Tsoory SG (2022) Inter-brain plasticity underlies empathic learning in social interactions.
Front. Psychol. 13:951248.
doi: 10.3389/fpsyg.2022.951248

COPYRIGHT
© 2022 Shamay-Tsoory. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Inter-brain plasticity underlies empathic learning in social interactions

Simone G. Shamay-Tsoory*

School of Psychological Sciences, University of Haifa, Haifa, Israel

KEYWORDS

empathy, inter-brain coupling, synchrony, learning, plasticity

Introduction

Empathy, our capacity to react to the suffering of others, is not a monolithic process and involves emotional (e.g., shared pain) and cognitive (e.g., perspective taking) components (Gonzalez-Liencres et al., 2013). While previous studies have focused on investigating the neural underpinnings of cognitive and emotional empathy in the target, it is increasingly acknowledged that integrative brain models for understanding the dynamic interaction between the target and the observer are warranted.

While the research on empathy emphasizes first and foremost its contribution to distress regulation in the target, few studies have examined how empathic responses of the observer actually change the state of the target. Indeed, although empathy occurs in social interactions, research on empathy have largely focused on covert mechanisms of empathy in the observer (empathizer), without exploring how empathic reactions affect the distress of the target (Main et al., 2017; Shamay-Tsoory and Mendelsohn, 2019).

To examine the role of empathy in regulating the target's distress, Reeck et al. (2016) have proposed a model of interpersonal emotion regulation that takes into account both the target and the observer. This model holds that empathy plays a major role in interpersonal emotion regulation, as the distress of the target may trigger an empathetic reaction in the observer. This model describes the participation of several empathy-related brain regions in the interpersonal emotion regulation cycle, including the dorsomedial prefrontal cortex (dmPFC), temporoparietal junction (TPJ) and inferior frontal gyrus (IFG). The dmPFC and TPJ are parts of the brains default mode network, a system that instantiates processes that support self-referential mental activity, mentalization and the recollection of prior experiences (Raichle, 2015). In addition to the default mode network, a central role in the empathy feedback loop is played by the observation-execution system also known as the mirror neurons system that includes the IFG as well as the inferior parietal lobe (IPL), regions which were suggested to play a role in emotional empathy (Shamay-Tsoory, 2011; Korisky et al., 2020). The mirror neurons, which were first discovered in the monkey ventral premotor cortex (area F5), discharge both during action performance and action observation (Rizzolatti et al., 1996; Rizzolatti and Craighero, 2004) and are believed to allow gaining knowledge of the observed action of others from a personal perspective (Buccino et al., 2004). Notably, it is increasingly acknowledged that the IFG is not only a structure that mediates speech production, but it is involved in action recognition (Buccino et al., 2004) and even in representing abstract

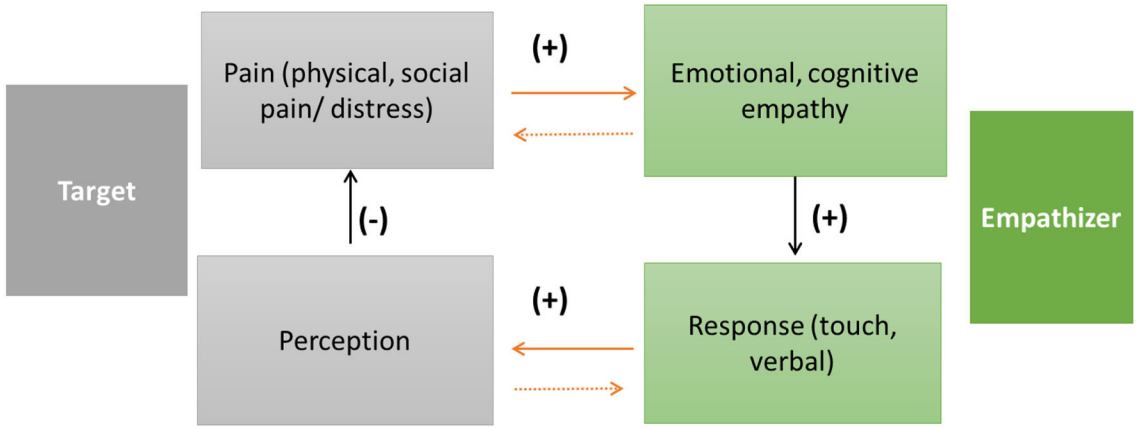


FIGURE 1
A model of inter-brain plasticity and empathy: the target is experiencing distress, which triggers empathy (emotional/cognitive) in the empathizer. Emotional and cognitive empathy contribute to reduction of distress by multiple means (e.g., mimicry, synchrony, and verbal responses). Activation in the observation-execution system is coupled between the empathizer and the target. This coupling reduces the target's distress by activating reward. As empathic interactions continue the empathizer learns how to adapt her reactions based on the target's feedback. As the empathizer adapts her response to the target, the inter-brain networks between them reconfigure.

representations of behavior (Del Maschio et al., 2022), which may allow representation of others goals intentions and emotions.

Although empathic interactions may unfold over time, the model of Reeck et al. (2016) does not address how empathic reactions change during interactions. To address this issue, Shamay-Tsoory and Hertz (2022) have recently coined the term *adaptive empathy*, to represent the ability to learn how to adapt one's responses to another's distress. The concept of adaptive empathy points out that it is essential to study how empathic reactions are adapted over time, based on feedback in the context of interactions between empathizer and target (Kozakevich-Arbel et al., 2021). In this context, the empathizer reacts to the distress of the target and may change their own response, based on feedback from the target. Examining how empathic reactions change over time represents a new approach, describing an empathic interactions feedback loop consisting of an empathizer providing responses that change during interactions, based on feedback from the target (see Figure 1).

The continuous updating of empathic responses demands the participation of a neural network that observes the target's actions and responses and activates the same representations of this behavior. Given the role of the IFG in action recognition (Buccino et al., 2004), which required continues updating of the others' behavior, this region may play a key role in empathic learning. Indeed, previous studies confirm that the IFG is essential for emotional empathy (Shamay-Tsoory et al., 2009) and that the IFG is activated during empathic learning (Hein et al., 2010), supporting the suggestion that this a core region in the adaptive empathy networks.

Notably, it was recently suggested that activations in the IFG may be demonstrated not only within a single brain, but also simultaneously recorded in the brains of interacting individuals (Shamay-Tsoory et al., 2019). Such inter-brain coupling represents coordinated brain activity of two or more interaction partners. Evidence from hyperscanning fNIRS studies shows that inter-brain coupling in the IFG of interacting dyads may underlie various forms of connection, from coordination during dialogues (Jiang et al., 2012) to movement synchrony (Gamliel et al., 2021) and singing in synchrony (Osaka et al., 2015). Furthermore, corroborating evidence from EEG studies further reveals that inter-brain coupling in the alpha band (8 to 12 or 13Hz), which is associated with the mirror neurons system, plays a role in empathic touch (Goldstein et al., 2018), suggesting that inter-brain coupling may also mediate affective empathy.

Given that empathic interaction develop over time, the question remains whether inter-brain coupling can increase over the course of one or multiple interactions. Recently it was suggested that *inter-brain plasticity*, the ability of interacting brains to modify the coupling between brains in reaction to repeated interactions underlies learning in social interactions (Shamay-Tsoory, 2021). The interbrain plasticity approach views the brain activity of interaction partners as components of an extended neural network that includes interbrain and intra-brain connections that change during interactions. In the case of empathy, it is possible that as the observer adapts her response to the target, the inter-brain networks between them reconfigure. In the initial phase of the interaction the observer may adapt her emotions to those of the target. This involves representing the behavior

of each other in the observation-execution system (Rizzolatti and Sinigaglia, 2016). The target observes her emotions mimicked by the observer, representing her emotions and then adapts her emotions to be aligned with the observer. This feedback is identified by the observer who may modify her emotions. During repeated interactions, the target and the observer represent each other's emotions in a similar manner and regions in their observation-execution system become gradually coupled. As inter-brain and intra-brain plasticity emerges, fewer sensorimotor signals are required to establish empathy. Over time the observer may improve her empathic responses and share their emotions better. This framework may explain how empathic responses may improve over time and how we learn to mutually adapt our responses during social intereactions.

Conclusion

An abundance of studies examined empathy by focusing on the empathizer, limiting our understanding of the interaction between the empathizer and the target during social interactions. Here, I integrate disparate lines of evidence into a new model of empathic learning. A feedback loop model of empathy is offered, one that accounts for learning how to change empathic reactions based on feedback over time. This model is supported by the new concept of inter-brain plasticity that examines changes in inter-brain coupling during interactions. While the literature on empathy discusses each of the stages of the model, no study to date has directly examined how brain-to-brain coupling change over time.

The model proposed here extends the interpersonal emotion regulation model of Reeck et al. (2016) by taking into account changes in the coupling between the observation execution systems of interaction partners over time. Changes in inter-brain coupling in the IFG represent a core component in this loop. This model offers new insight on the neural basis of empathy and may have clinical implications for understanding population with empathy difficulties.

Author contributions

The author confirms being the sole contributor of this work and has approved it for publication.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Buccino, G., Binkofski, F., and Riggio, L. (2004). The mirror neuron system and action recognition. *Brain Lang.* 89, 370–376. doi: 10.1016/S0093-934X(03)00356-0

Del Maschio, N., Fedeli, D., Garofalo, G., and Buccino, G. (2022). Evidence for the concreteness of abstract language: a meta-analysis of neuroimaging studies. *Brain Sci.* 12, 35–37. doi: 10.3390/brainsci12010032

Gamliel, H. N., Nevat, M., Probolovski, H. Z. G., Karklinsky, M., Han, S., Shamay-Tsoory, S. G., et al. (2021). Inter-group conflict affects inter-brain synchrony during synchronized movements. *NeuroImage* 245, 118661. doi: 10.1016/j.neuroimage.2021.118661

Goldstein, P., Weissman-Fogel, I., Dumas, G., and Shamay-Tsoory, S. G. (2018). Brain-to-brain coupling during handholding is associated with pain reduction. *Proc. Nat. Acad. Sci. U.S.A* 115, E2528–E2537. doi: 10.1073/pnas.1703643115

Gonzalez-Liencres, C., Shamay-Tsoory, S. G., and Brüne, M. (2013). Towards a neuroscience of empathy: Ontogeny, phylogeny, brain mechanisms, context and psychopathology. *Neurosci. Biobehav. Rev.* 37, 1537–1548. doi: 10.1016/j.neubiorev.2013.05.001

Hein, G., Silani, G., Preuschoff, K., Batson, C. D., and Singer, T. (2010). Neural responses to ingroup and outgroup members' suffering predict individual differences in costly helping. *Neuron* 68, 149–160. doi: 10.1016/j.neuron.2010.09.003

Jiang, J., Dai, B., Peng, D., Zhu, C., Liu, L., Lu, C., et al. (2012). Neural synchronization during face-to-face communication. *J. Neurosci.* 32, 16064–16069. doi: 10.1523/JNEUROSCI.2926-12.2012

Korisky, A., Eisenberger, N. I., Nevat, M., Weissman-Fogel, I., and Shamay-Tsoory, S. G. (2020). A dual-brain approach for understanding the neural mechanisms that underlie the comforting effects of social touch. *Cortex* 127, 333–346. doi: 10.1016/j.cortex.2020.01.028

Kozakevich-Arbel, E., Shamay-Tsoory, S., and Hertz, U. (2021). Adaptive Empathy: Empathic response selection as a dynamic, feedback-based learning process. *Front. Psychiatry* 12, 1248. doi: 10.3389/fpsyt.2021.706474

Main, A., Walle, E. A., Kho, C., and Halpern, J. (2017). The interpersonal functions of empathy: a relational perspective. *Emot. Rev.* 9, 358–366. doi: 10.1177/1754073916669440

Osaka, N., Minamoto, T., Yaoi, K., Azuma, M., Shimada, Y. M., Osaka, M., et al. (2015). How two brains make one synchronized mind in the inferior frontal cortex: FNIRS-based hyperscanning during cooperative singing. *Front. Psychol.* 6, 1811. doi: 10.3389/fpsyg.2015.01811

Raichle, M. E. (2015). The brain's default mode network. *Ann. Rev. Neurosci.* 38, 433–447. doi: 10.1146/annurev-neuro-071013-014030

Reeck, C., Ames, D. R., and Ochsner, K. N. (2016). The social regulation of emotion: an integrative, cross-disciplinary model. *Trends Cognit. Sci.* 20, 47–63. doi: 10.1016/j.tics.2015.09.003

Rizzolatti, G., and Craighero, L. (2004). The mirror-neuron system. *Ann. Rev. Neurosci.* 27, 169–192. doi: 10.1146/annurev.neuro.27.070203.144230

Rizzolatti, G., Fadiga, L., Gallese, V., and Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognit. Brain Res.* 3, 131–141. doi: 10.1016/0926-6410(95)00038-0

Rizzolatti, G., and Sinigaglia, C. (2016). The mirror mechanism: a basic principle of brain function. *Nat. Rev. Neurosci.* 17, 757–765. doi: 10.1038/nrn.2016.135

Shamay-Tsoory, S. G. (2011). The neural bases for empathy. *Neuroscientist* 17, 18–24. doi: 10.1177/1073858410379268

Shamay-Tsoory, S. G. (2021). Brains that fire together wire together: interbrain plasticity underlies learning in social interactions. *Neuroscientist* 1073858421996682. doi: 10.1177/1073858421996682

Shamay-Tsoory, S. G., Aharon-Peretz, J., and Perry, D. (2009). Two systems for empathy: a double dissociation between emotional and cognitive empathy in inferior frontal gyrus versus ventromedial prefrontal lesions. *Brain* 132, 617–627. doi: 10.1093/brain/awn279

Shamay-Tsoory, S. G., and Hertz, U. (2022). Adaptive empathy: a model for learning empathic responses in response to feedback. *Pers. Psychol. Sci.* 17, 1008–1023. doi: 10.1177/17456916211031926

Shamay-Tsoory, S. G., and Mendelsohn, A. (2019). Real-life neuroscience: an ecological approach to brain and behavior research. *Pers. Psychol. Sci.* 14, 841–859. doi: 10.1177/1745691619856350

Shamay-Tsoory, S. G., Saporta, N., Marton-Alper, I. Z., and Gvirts, H. Z. (2019). Herding brains: a core neural mechanism for social alignment. *Trends Cognit. Sci.* 23, 174–186. doi: 10.1016/j.tics.2019.01.002

frontiers | Frontiers in Psychology

TYPE Opinion
PUBLISHED 22 July 2022
DOI 10.3389/fpsyg.2022.943765

Check for updates

OPEN ACCESS

EDITED BY
Rüdiger J. Seitz,
Heinrich Heine University of Düsseldorf, Germany

REVIEWED BY
Loïc P. Heurley,
Université Paris Nanterre, France

*CORRESPONDENCE
Giovanni Buccino
buccino.giovanni@hsr.it

SPECIALTY SECTION
This article was submitted to Cognition, a section of the journal Frontiers in Psychology

RECEIVED 14 May 2022
ACCEPTED 01 July 2022
PUBLISHED 22 July 2022

CITATION
Buccino G and Colagè I (2022) Grounding abstract concepts and beliefs into experience: The embodied perspective.
Front. Psychol. 13:943765.
doi: 10.3389/fpsyg.2022.943765

COPYRIGHT
© 2022 Buccino and Colagè. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Grounding abstract concepts and beliefs into experience: The embodied perspective

Giovanni Buccino[1]* and Ivan Colagè[2,3]

[1]Divisione di Neuroscienze, IRCCS San Raffaele and Università Vita-Salute San Raffaele, Milan, Italy, [2]Faculty of Philosophy, Pontifical University Antonianum, Rome, Italy, [3]DISF Research Center, Pontifical University of the Holy Cross, Rome, Italy

KEYWORDS

embodied language, abstract concepts, beliefs, semantic hub, Broca's region

Introduction

The embodied perspective on language is now supported by several studies showing that activation of neural substrates processing the sensory and motor aspects of the world is not only associated with the processing of language referring to concrete aspects of the world (Buccino et al., 2005; Tettamanti et al., 2005; Gough et al., 2012, 2013; Marino et al., 2013, 2014; Visani et al., 2022) but is also *causal* to the understanding of concrete language (e.g., Bak et al., 2006; Sato et al., 2008; Kemmerer et al., 2012; Tremblay et al., 2012; Cardona et al., 2013; Fernandino et al., 2013; Desai et al., 2015; Klepp et al., 2015; Buccino et al., 2018; for review Buccino et al., 2016).

Less clear is the situation about abstract domains – i.e., language items referring to less concrete actions (e.g., "I give you my *opinion*") or less tangible aspects of the world (e.g., "freedom") (Glenberg et al., 2008; Boulenger et al., 2009). Such language items bring in a more conceptual and lexical-semantic dimension apparently less amenable to be understood in embodied terms. However, such abstract and conceptual domains are widespread in human linguistic practice and are, consequently, both problematic and interesting for the embodied approach (Gallese and Lakoff, 2005). Limited empirical findings and the variety of theoretical stances on abstract language (see Binder et al., 2009; Wang et al., 2010; Buccino et al., 2019 for reviews) have prompted "hybrid" models on abstract concepts and words. All these models share the embodied approach, but all also posit that acquisition and understanding of abstract concepts and words is only partially grounded in experience-related sensory-motor neural substrates and also resorts to supposed a-modal brain modules processing "pure" language aspects.

In the next section we will first briefly mention such hybrid models; then we will present a "fully embodied" approach (Buccino et al., 2019) and review the available evidence supporting it (Del Maschio et al., 2021). Finally, we will suggest how the advancements in embodied abstract language may shed light on the nature of beliefs.

From embodied abstract language to beliefs

The hybrid models for abstract language and concepts

Three main hybrid models for embodied abstract language have been proposed. The first model (Borghi et al., 2017, 2018) forwards that abstract words and concepts are mainly rooted in the social conventions and ensuing social interactions about abstract content, thus implying the existence of a brain system dedicated to processing propositional aspects of language. The second model maintains that specific features of words' meanings are indeed coded in sensory, motor or even emotional brain circuits. However, words' meaning is ultimately coded in specific, high-order, a-modal, linguistic regions, (Binder et al., 2009; Desai et al., 2015; Mahon, 2015) labeled as "semantic hubs". The third model views the specificity of abstract words and concepts in the exalted emotional load they display and thus forwards that processing abstract contents specifically involves brain regions for coding, feeling and expressing emotions (Barsalou, 1999; Kousta et al., 2011; Moseley et al., 2012; Vigliocco et al., 2014). It is worth stressing here, however, that a number of studies (Wilson-Mendenhall et al., 2011, 2013) have shown that emotions themselves are grounded in the neural structures where the experiences and experiential contexts emotion word refer to are represented.

A fully embodied approach

Buccino et al. (2019) forwarded a fully embodied approach to abstract language, that avoids assuming the existence of a-modal, purely linguistic systems to processing abstract words and concepts. This proposal is based on the idea that abstract words and concepts are such because of the complexity of the experiences attached to them, and not because they are far or detached from concrete experience. Specifically, such experiential complexity can increase according to (i) the number of effectors involved, (ii) the number of sensory systems engaged, as well as (iii) the accumulation over time of concrete life experiences (and related emotional load) attached to those words/concepts. Moreover, the distinction between abstract and concrete words/concepts may be one of degree and not of kind, as the complexity of experiences may increase along a continuum rather than sharply.

This approach allows for a strongly embodied interpretation of the evidence about the neural substrates processing abstract words, thus overcoming the need to elaborate hybrid models. Besides the data reviewed to advocate this fully embodied approach (see also Buccino et al., 2016, 2019), a recent meta-analysis of neuroimaging studies reporting activations related to abstract and concrete concepts further support this fully embodied approach (Del Maschio et al., 2021).

This meta-analysis shows that extensive clusters in the left temporal lobe (including the middle and inferior temporal gyri) and in the left motor cortex, as well as activations in right parietal cortex, left inferior frontal gyrus, and prefrontal regions are found for *both* concrete and abstract concepts. This suggests that (a) processing of these two kinds of concepts is not sharply segregated in the brain, (b) abstract concepts, like concrete ones, engage brain circuits involved in subjects' interaction with the world, and (c) abstract concepts are not pre-eminently processed in linguistic/propositional format, in semantic hubs or in emotion-related areas (in contrast to what hybrid models propose). Consequently, since semantic hubs are neural structures engaged by both concrete and abstract concepts, it is hard to accept the notion that they may be the "apex" of hierarchical structures progressively moving from processing concrete to abstract situations. Rather, these semantic hubs may play the role to contextualize actions (and related linguistic material) independently of their degree of abstractness.

The metanalysis by Del Maschio et al. (2021) also unveils that brain regions more active for abstract than concrete concepts encompass two major clusters in the left inferior frontal gyrus (pars triangularis and orbitalis, largely overlapping Broca's region) and middle temporal gyrus, as well as smaller clusters in medial frontal cortex and bilateral temporal poles.

According to the hybrid models, the stronger activation of Broca's region during the processing of abstract language supports the notion that abstract language is coded in a propositional format, since Broca's region is classically considered a linguistic region. In contrast with the classical view, many functions are now attributed to the Broca's region (Amunts and Zilles, 2012; Hardwick et al., 2018). First, in Broca's region there is a motor representation of mouth, hand-arm and, likely, foot actions (Binkofski et al., 1999; Nishitani et al., 2005). Secondly, Broca's region also processes observed and imagined actions (Binder et al., 2009; Hardwick et al., 2018). Thirdly, and more generally, there is also representation of mimicked actions, i.e., actions where the effector is used independently of the object (Lui et al., 2008); mimicked actions may be regarded as a first step in generalizing over object-oriented actions. Finally, Broca's region also codes actions able to mediate a semantic meaning, such as in emblems, but always using a biological effector (Andric et al., 2013).

All this suggests that Broca's region can support a process of generalization (indeed, of abstraction) of actions, but always starting from concrete situations and contexts: it might be said that Broca's region can grasp "what is common" to various instantiations of actions in varying contexts and situations. This view of Broca's activation is consistent with the notion that abstract language engages multiple effectors and contexts in which the use of the effector is not bound to specific objects. Put differently, because abstract concepts and their

corresponding verbal labels express actions or entities that are dynamic in time and space, executed by different effectors, and coded in different systems (Buccino et al., 2019), their content, more strongly than for concrete concepts and words, is coded "motorically" in a brain region where actions are represented in a conceptual manner.

Other areas found more active for abstract language (specifically, medial frontal cortex and middle temporal gyrus) are indeed part of the proposed a-modal semantic hubs (Binder et al., 2009; Desai et al., 2015; Mahon, 2015). However, these brain regions are also known to be part of the "default-mode" network that is modulated by demanding cognitive tasks or by social cognition (Mars et al., 2012; Raichle, 2015); their engagement in processing abstract language (i.e., language items attached to complex experience) can be explained assuming that they may contribute to define an appropriate context for the processed words and their link with life experiences and personal beliefs.

Summing up, a fully embodied approach would account for the available data about processing of abstract language in the brain consistently with current knowledge of the functions of brain regions not directly involved in sensory-motor processing and without postulating the existence of a-modal, purely linguistic brain modules.

Implications for belief

Beliefs are high mental processes implying abstract conceptualization and generalization. In this context, the notion of "belief" should be understood broadly, so to encompass moral contexts related to value and religion as well as cognitive convictions on how the world is done and works (and on how we should consequently behave in it). Moreover, beliefs should be conceptualized in strict connection with actions and life conduct. Philosopher C. Peirce stated that "a conception, that is, the rational purport of a word or other expression lies exclusively in its conceivable bearing upon the conduct of life" (Peirce, 1905; p. 162). A recent neuroscientific approach to beliefs indeed posits that selection of beliefs is virtually equivalent to selection of actions (Sugiura et al., 2015). The link between beliefs and actions, as well as the understanding of beliefs as conceptual items, suggests the relevance of an embodied approach to abstractness for the issue of beliefs.

Interestingly, recent developments in the neuroscience of action has established a link between action-related brain processes and the issue of beliefs and personal identity (Jeannerod, 2001, 2006, 2009; see also Colagè and Gobbi, 2017). According to this theory, the assessment of our experiences, especially the results of complex actions, may lead us to build up, and possibly revise, our belief system, which in turn allows for the planning of complex actions (Jeannerod, 2009, p. 263–269).

For this reason, a fully embodied approach to abstract words and concepts may shed light onto the process of building up and revising beliefs, specifically suggesting that beliefs, much like other conceptual domains, can be grounded in actual experiences and their complexity. Three further hints can be added.

First, we have seen that mesial pre-frontal cortex activates in processing abstract language and that this activation can be explained by the need to contextualize and frame abstract words on the background of one's life experiences. Specifically, studies suggest that this brain region is modulated, during the judgment of different relevant social situations and contexts, by the degree of similarity with our own beliefs, attitudes and inclinations (Mitchell et al., 2006; Zaki et al., 2014). Given the complexity of experiences attached to abstract words, these activations may help focusing on a relevant subset of the complex array of experiences attached to an abstract word. It is interesting to note that mesial pre-frontal cortex is proposed as key structure for processing beliefs, and specifically for integrating perception-, action- and emotion/value-related information (Seitz and Angel, 2012; Sugiura et al., 2015).

Secondly, mesial pre-frontal cortices are also known to be part of the mentalizing and affect-related brain systems (Frith and Frith, 2012). Activation of mesial frontal cortex in processing abstract words/concepts may reflect the need of the subject to retrieve his/her social and self-related beliefs to understand abstract linguistic items properly (see also Buccino and Colagè, 2017).

Finally, a fully embodied approach to beliefs is also consistent with the idea that such linguistic transactions among human beings are anyway grounded in real experiences. Linguistic transactions are effective in belief formation to the extent to which they help us sharing and combining our experiential baggage (Colagè and Buccino, 2016).

Author contributions

GB conceptualized the paper and revised the manuscript. IC wrote the first draft of the manuscript and revised it. All authors contributed to the article and approved the submitted version.

Funding

This paper is funded by Dr. Rüdiger Seitz, *via* the Volkswagen Foundation, Siemens Healthineers, and the Betz Foundation.

Acknowledgments

Both authors are grateful to the Special Issue Editors, and especially to H.-F. Angel and R. Setiz for past meetings they

organized in Graz and the unique opportunity they offered to share ideas and experiences in the field of beliefs and credition.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Amunts, K., and Zilles, K. (2012). Architecture and organizational principles of Broca's region. *Trends Cogn. Sci.* 16, 418–426. doi: 10.1016/j.tics.2012.06.005

Andric, M., Solodkin, A., Buccino, G., Goldin-Meadow, S., and Rizzolatti, G., Small, S.L. (2013). Brain function overlaps when people observe emblems, speech, and grasping. *Neuropsychologia* 51, 1619–1629. doi: 10.1016/j.neuropsychologia.2013.03.022

Bak, T. H., Yancopoulou, D., Nestor, P. J., Xuereb, J. H., Spillantini, M. G., Pulvermüller, F., et al. (2006). Clinical, imaging and pathological correlates of a hereditary deficit in verb and action processing. *Brain* 129, 321–332. doi: 10.1093/brain/awh701

Barsalou, L. W. (1999). Perceptual symbol systems and emotion. *Behav. Brain Sci*. 22, 612–613. doi: 10.1017/S0140525X99252144

Binder, J. R., Desai, R. H., Graves, W. W., and Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb. Cortex* 19, 2767–2796. doi: 10.1093/cercor/bhp055

Binkofski, F., Buccino, G., Posse, S., Seitz, R. J., Rizzolatti, G., and Freund, H. J. (1999). A fronto-parietal circuit for object manipulation in man: Evidence from an fMRI-study. *Eur. J. Neurosci.* 11, 3276–3286. doi: 10.1046/j.1460-9568.1999.00753.x

Borghi, A. M., Barca, L., Binkofski, F., and Tummolini, L. (2018). Abstract concepts, language and sociality: from acquisition to inner speech. *Philos. Trans. R. Soc. B Biol. Sci.* 373, 20170134. doi: 10.1098/rstb.2017.0134

Borghi, A. M., Binkofski, F., Castelfranchi, C., Cimatti, F., Scorolli, C., and Tummolini, L. (2017). The challenge of abstract concepts. *Psychol. Bull.* 143, 263–292. doi: 10.1037/bul0000089

Boulenger, V., Hauk, O., and Pulvermueller, F. (2009). Grasping ideas with the motor system: Semantic somatotopy in idiom comprehension. *Cereb. Cortex* 19, 1905–1914. doi: 10.1093/cercor/bhn217

Buccino, G., and Colagè, I. (2017). "New concepts on the motor system: implications for emotions and creditions", in *Processes of Believing: The Acquisition, Maintenance, and Change in Creditions*, eds H. -F.Angel, R. F. Paloutzian, A. L. C. Runehov, and R. J. Seitz (Cham: Springer International Publishing AG), 97–109. doi: 10.1007/978-3-319-50924-2_7

Buccino, G., Colagè, I., Gobbi, N. and, Bonaccorso, G. (2016). Grounding meaning in experience: a broad perspective on embodied language. *Neurosci. Biobehav. Rev*. 69, 69–78. doi: 10.1016/j.neubiorev.2016.07.033

Buccino, G., Colagè, I., Silipo, F. and, D'Ambrosio, P. (2019). The concreteness of abstract language: an ancient issue and a new perspective. *Brain Struct. Funct.* 224, 1385–1401. doi: 10.1007/s00429-019-01851-7

Buccino, G., Dalla Volta, R., Arabia, G., Morelli, M., Chiriaco, C., Lupo, A., et al. (2018). Processing graspable object images and their nouns is impaired in Parkinson's disease patients. *Cortex* 100, 32–39. doi: 10.1016/j.cortex.2017.03.009

Buccino, G., Riggio, L., Melli, G., Binkofski, F., Gallese, V., and Rizzolatti, G. (2005). Listening to action-related sentences modulates the activity of the motor system: a combined TMS and behavioral study. *Cogn. Brain Res.* 24, 355–363. doi: 10.1016/j.cogbrainres.2005.02.020

Cardona, J. F., Gershanik, G., Gelormini-Lezama, C., Houck, A. L., Cardona, S., Kargieman, L., et al. (2013). Action-verb processing in Parkinson's disease: new pathways for motor-language coupling. *Brain Struct. Funct.* 218, 1355–1373. doi: 10.1007/s00429-013-0510-1

Colagè, I., and Buccino, G. (2016). Il linguaggio e l'esperienza: dalle neuroscienze all'ambiente naturale e culturale. *Antonianum* 91, 1047–1061.

Colagè, I., and Gobbi, N. (2017). "Creditions and Self-Identity: Embodiment Between Actions and (Life) Decisions", in Processes of Believing: The Acquisition, Maintenance, and Change in Creditions, eds H. -F.Angel, R. F. Paloutzian, A. L. C. Runehov, and R. J. Seitz (Springer International Publishing AG), 247–262. doi: 10.1007/978-3-319-50924-2_18

Del Maschio, N., Fedeli, D., Garofalo, G., and Buccino, G. (2021). Evidence for the concreteness of abstract language: a meta-analysis of neuroimaging studies. *Brain Sci.* 12, 32. doi: 10.3390/brainsci12010032

Desai, R. H., Herter, T., Riccardi, N., Rorden, C., and Fridriksson, J. (2015). Concepts within reach: action performance predicts action language processing in stroke. *Neuropsychologia* 71, 217–224. doi: 10.1016/j.neuropsychologia.2015.04.006

Fernandino, L., Conant, L. L., Binder, J. R., Blindauer, K., Hiner, B., Spangler, K., et al. (2013). Where is the action? Action sentence processing in Parkinson's disease. *Neuropsychologia* 51, 1510–1517. doi: 10.1016/j.neuropsychologia.2013.04.008

Frith, C. D., and Frith, U. (2012). Mechanisms of social cognition. *Annu. Rev. Psychol.* 63, 287–313. doi: 10.1146/annurev-psych-120710-100449

Gallese, V., and Lakoff, G. (2005). The brain's concepts: the role of the sensory-motor system in conceptual knowledge. *Cognit. Neuropsychol.* 22, 455–479. doi: 10.1080/02643290442000310

Glenberg, A. M., Sato, M., Cattaneo, L., Riggio, L., Palumbo, D., and Buccino, G. (2008). Processing abstract language modulates motor system activity. *Q. J. Exp. Psychol.* 61 905–919. doi: 10.1080/17470210701625550

Gough, P. M., Campione, G. C., and Buccino, G. (2013). Fine tuned modulation of the motor system by adjectives expressing positive and negative properties. *Brain Lang.* 125, 54–59. doi: 10.1016/j.bandl.2013.01.012

Gough, P. M., Riggio, L., Chersi, F., Sato, M., Fogassi, L., and Buccino, G. (2012). Nouns referring to tools and natural objects differentially modulate the motor system. *Neuropsychologia* 50, 19–25. doi: 10.1016/j.neuropsychologia.2011.10.017

Hardwick, R. M., Caspers, S., Eickhoff, S. B., and Swinnen, S. P. (2018). Neural correlates of action: comparing meta-analyses of imagery, observation, and execution. *Neurosci. Biobehav. Rev.* 94, 31–44. doi: 10.1016/j.neubiorev.2018.08.003

Jeannerod, M. (2001). Neural simulation of actions: a unifying mechanism for motor cognition. *NeuroImage* 14, 103–109. doi: 10.1006/nimg.2001.0832

Jeannerod, M. (2006). *Motor Cognition: What Actions Tell the Self*. Oxford: Oxford University Press.

Jeannerod, M. (2009). *Le Cerveau Volontaire*. Paris: Odile Jacob.

Kemmerer, D., Rudrauf, D., Manzel, K., and Tranel, D. (2012). Behavioral patterns and lesion sites associated with impaired processing of lexical and conceptual knowledge of actions. *Cortex* 48, 826–848. doi: 10.1016/j.cortex.2010.11.001

Klepp, A., Niccolai, V., Buccino, G., Schnitzler, A., and Biermann-Rubena, K. (2015). Language–motor interference reflected in MEG beta oscillations. *NeuroImage* 109, 438–448. doi: 10.1016/j.neuroimage.2014.12.077

Kousta, S.-T., Vigliocco, G., Vinson, D. P., Andrews, M., and Del Campo, E. (2011). The representation of abstract words: why emotion matters. *J. Exp. Psychol. Gen.* 140, 14–34. doi: 10.1037/a0021446

Lui, F., Buccino, G., Duzzi, D., Benuzzi, F., Crisi, G., Baraldi, P., et al. (2008). Neural substrates for observing and imagining non-object-directed actions. *Soc. Neurosci.* 3, 261–275. doi: 10.1080/17470910701458551

Mahon, B. Z. (2015). What is embodied about cognition? *Lang. Cogn. Neurosci.* 30, 420–429. doi: 10.1080/23273798.2014.987791

Marino, B., Sirianni, M., Dalla Volta, R., Magliocco, F., Silipo, F., Quattrone, A., et al. (2014). Viewing photos and reading nouns of natural graspable objects similarly modulate motor responses. *Front. Hum. Neurosci.* 8, 968. doi: 10.3389/fnhum.2014.00968

Marino, B. F., Gough, P. M., Gallese, V., Riggio, L., and Buccino, G. (2013). How the motor system handles nouns: a behavioral study. *Psychol. Res.* 77, 64–73. doi: 10.1007/s00426-011-0371-2

Mars, R. B., Neubert, F. X., Noonan, M. A. P., Sallet, J., Toni, I., and Rushworth, M. F. S. (2012). On the relationship between the "default mode network" and the "social brain". *Front. Hum. Neurosci.* 6, 189. doi: 10.3389/fnhum.2012.00189

Mitchell, J. P., Macrae, C. N., and Banaji, M. R. (2006). Dissociable medial prefrontal contribution to judgements of similar and dissimilar others. *Neuron* 18, 655–663. doi: 10.1016/j.neuron.2006.03.040

Moseley, R., Carota, F., Hauk, O., Mohr, B., and Pulvermüller, F. (2012). A role for the motor system in binding abstract emotional meaning. *Cereb. Cortex* 22, 1634–1647. doi: 10.1093/cercor/bhr238

Nishitani, N., Schürmann, M., Amunts, K. and, Hari, R. (2005). Broca's region: from action to language. *Physiology* 20, 60–69. doi: 10.1152/physiol.00043.2004

Peirce, C. S. (1905). What pragmatism is. *Monist* 15, 161–181. doi: 10.5840/monist190515230

Raichle, M. E. (2015). The brain's default mode network. *Annu. Rev. Neurosci.* 38, 433–447. doi: 10.1146/annurev-neuro-071013-014030

Sato, M., Mengarelli, M., Riggio, L., Gallese, V., and Buccino, G. (2008). Task related modulation of the motor system during language processing. *Brain Lang.* 105, 83–90. doi: 10.1016/j.bandl.2007.10.001

Seitz, R. J., and Angel, H. F. (2012). Processes of believing – a review and conceptual account. *Rev Neurosci.* 23, 303–309. doi: 10.1515/revneuro-2012-0034

Sugiura, M., Seitz, R. J., and Angel, H. F. (2015). Models and neural bases of the believing process. *J. Behav. Brain. Sci.* 5, 12–23. doi: 10.4236/jbbs.2015.51002

Tettamanti, M., Buccino, G., Saccuman, M. C., Gallese, V., Danna, M., Scifo, P., et al. (2005). Listening to action-related sentences activates fronto-parietal motor circuits. *J. Cogn. Neurosci.* 17, 273–281. doi: 10.1162/0898929053124965

Tremblay, P., Sato, M., and Small, S. L. (2012). TMS-induced modulation of action sentence priming in the ventral premotor cortex. *Neuropsychologia* 50, 319–326. doi: 10.1016/j.neuropsychologia.2011.12.002

Vigliocco, G., Kousta, S.-T., Della Rosa, P. A., Vinson, D. P., Tettamanti, M., Devlin, J. T., et al. (2014). The neural representation of abstract words: the role of emotion. *Cereb. Cortex* 24, 1767–1777. doi: 10.1093/cercor/bht025

Visani, E., Rossi Sebastiano, D., Duran, D., Garofalo, G., Magliocco, F., Silipo, F., et al. (2022). The semantics of natural objects and tools in the brain: a combined behavioral and MEG study. *Brain Sci.* 12, 97. doi: 10.3390/brainsci12010097

Wang, J., Conder, J. A., Blitzer, D. N., and Shinkareva, S. V. (2010). Neural representation of abstract and concrete concepts: a meta-analysis of neuroimaging studies. *Hum. Brain Mapp.* 31, 1459–1468. doi: 10.1002/hbm.20950

Wilson-Mendenhall, C. D., Barrett, L. F., and Barsalou, L. W. (2013). Situating emotional experience. *Front. Human Neurosci.* 26, 764. doi: 10.3389/fnhum.2013.00764

Wilson-Mendenhall, C. D., Barrett, L. F., Simmons, W. K., and Barsalou, L. W. (2011). Grounding emotion in situated conceptualization. *Neuropsychologia* 49, 1105–1127. doi: 10.1016/j.neuropsychologia.2010.12.032

Zaki, J., Lòpez, G., and Mitchell, J. P. (2014). Activity in ventromedial prefrontal cortex co-varies with revealed social preferences: evidence for person-invariant value. *Soc. Cogn. Affect Neurosci.* 9, 464–469. doi: 10.1093/scan/nst005

frontiers | Frontiers in Behavioral Neuroscience

TYPE Opinion
PUBLISHED 22 July 2022
DOI 10.3389/fnbeh.2022.962225

Check for updates

OPEN ACCESS

EDITED BY
Rüdiger J. Seitz,
Heinrich Heine University
of Düsseldorf, Germany

REVIEWED BY
Nina Dalkner,
Medical University of Graz, Austria

*CORRESPONDENCE
Shihui Han
shan@pku.edu.cn

SPECIALTY SECTION
This article was submitted to
Emotion Regulation and Processing,
a section of the journal
Frontiers in Behavioral Neuroscience

RECEIVED 06 June 2022
ACCEPTED 28 June 2022
PUBLISHED 22 July 2022

CITATION
Han S (2022) Cultural differences in
beliefs and believing about the self –
A brain imaging approach.
Front. Behav. Neurosci. 16:962225.
doi: 10.3389/fnbeh.2022.962225

COPYRIGHT
© 2022 Han. This is an open-access
article distributed under the terms of
the Creative Commons Attribution
License (CC BY). The use, distribution
or reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Cultural differences in beliefs and believing about the self – A brain imaging approach

Shihui Han*

Beijing Key Laboratory of Behavior and Mental Health, School of Psychological and Cognitive Sciences, PKU-IDG/McGovern Institute for Brain Research, Peking University, Beijing, China

KEYWORDS

culture, belief, believe, fMRI, medial prefrontal cortex

Introduction

Beliefs refer to "fundamental representations of imaginative and emotional content that link an individual's prior experience with his/her future behavior" (Seitz, 2022). We have beliefs about our societies, other people, and ourselves. When I say "I believe I will be able to contribute an article to this book by working with my colleagues," I express beliefs about myself and other people that help my decision-making regarding actions in the future. However, due to differences in social information that individuals receive and differences in sociocultural environments in which individuals develop and live, beliefs vary greatly across different human societies. Distinct beliefs not only influence people's behaviors but also shape how their brains work. In this study, I sought to briefly summarize previous research on cultural differences in beliefs about the self and relevant neural underpinnings. I further introduce a recent brain imaging approach to the neural correlate of believing and its potential cultural differences. Finally, I discuss the implications of the transcultural brain imaging findings related to belief and believing.

Cultural differences in beliefs

What is the nature of the self as a social unit in human societies? Representations of the self are built based on individuals' prior experiences and guide their social/economic/political decision-making. Understanding beliefs about the self has been the goal of research in multiple disciplines including philosophy, psychology, and neuroscience. Interestingly, beliefs regarding the self vary tremendously across different populations and different societies and have distinct neural correlates. For example, Western thoughts regard the self as a particular being that is distinct from others (Seigel, 2005) whereas Chinese thoughts take the self as a knot in a social network that unifies numerous individuals as a whole (Zhang, 2005).

These philosophical thoughts are formulated in psychology by Markus and Kitayama (1991) who claimed that Western (European and North American in particular) cultures teach people to view the self as an independent and autonomous entity that is inclined to attend to the self-more than others and emphasizes unique

dispositions or traits of the self (an independent view of the self). East Asian cultures, however, view the self that is fundamentally connected with others and thus are sensitive to information about significant others. The unique cultural beliefs about the self are tested empirically in both behavioral and brain imaging studies. It has been shown that Westerners remember self-related information better than information about close others such as mother and best friend (Klein et al., 1989; Heatherton et al., 2006), whereas Chinese remember information about self and close others equally well (Zhu and Zhang, 2002). These findings suggest distinct cultural beliefs about the self and are supported by transcultural brain imaging research on neural underpinnings of representations of the self.

An early functional magnetic resonance imaging (fMRI) study scanned both Chinese and Western students during judgments on personality traits of the self and a close other (i.e., mother) (Zhu et al., 2007). The results revealed activations in the ventral medial prefrontal cortex (mPFC) in response to the reflection of one's own traits in both Chinese and Western students. However, reflections of the self and mother evoked overlapping mPFC activity in Chinese, whereas Westerners showed greater mPFC activity during reflection of the self (vs. mother). These fMRI findings implicate overlapping neural representations of beliefs about the self and a close other in Chinese but not in Western students. A following fMRI study scanned both Chinese and Danish students, using fMRI, while the participants reflected personality traits, physical features, or social roles of themselves or a familiar celebrity in their own country (Ma et al., 2014). It was found that, although both Chinese and Danish students showed increased ventral mPFC activity during self-reflection of the self (vs. celebrity), the ventral mPFC activity was greater in Danish than in Chinese students. However, self-reflection on social roles in Chinese but not Danish students activated the temporoparietal junction, which is engaged in the processing of others' minds (Saxe and Kanwisher, 2003). In addition, Chinese compared to Danish students showed stronger functional connectivity between the ventral mPFC and the TPJ associated with self-reflection on social attributes.

Beliefs about the self also differ substantially between religious believers and non-believers. An fMRI study of non-religious and Christian Chinese found that, while the non-religious participants showed activations in the ventral mPFC during reflection of the self, self-reflection activated the dorsal mPFC in Christians (Han et al., 2008). Because the dorsal mPFC is usually involved in the inference of others' mental states (Grèzes et al., 2004), it was speculated that the dorsal mPFC supports beliefs regarding the self that underscores the evaluative process of the self by God. An fMRI study of Chinese Buddhists showed that reflection of the self-activated the dorsal mPFC, the rostral anterior cingulate and midcingulate cortices, and the left frontal/insular cortices (Han et al., 2010). It is likely that, during the self-reflection task, Buddhists might have to monitor the conflict between the doctrine of no-self and self-focus thinking during self-trait judgment. Together, these brain imaging findings suggest that, in response to cultural group differences in beliefs regarding the self, the human brain has evolved distinct patterns of neural activities that support specific beliefs about the self.

Cultural differences in believing

Early brain studies focused on cultural differences in neural representations of beliefs regarding the self. Little research examined neural correlates of believing as a process. Conceptually, believing may include multiple mental operations that support perception, valuation, information storage, and prediction (Angel and Seitz, 2016). Methodologically, it is a challenge to disentangle the neurocognitive processes of believing from other mental processes by controlling perceptual/cognitive/affective processes that do not essentially characterize believing. Based on the assumption that the believing process is connected with personal relevance and deals with a set of knowledge with a hierarchically organized structure (Sugiura et al., 2015), we designed a believing task and a control task in an fMRI study to disentangle neural processes of believing (Han et al., 2017). "Believe" and "think" are two words that are used to mutually explain each other in lay opinions (Allen et al., 1990) and may consist of overlapping mental processes. While being presented with trait adjectives during fMRI scanning, one group of Chinese participants was asked to make a yes or no response to the question "Do you *believe* that the trait adjective describes you (or a celebrity)?" (Belief group). Another group of Chinese participants made a yes or no response to the question "Do you *think* the trait adjective describes you (or a celebrity)?" (Think group). The same set of trait adjectives was used in different judgment tasks and judgments on a celebrity were employed as a control condition. We examined believing specific neurocognitive processes by comparing brain activities of Believe vs. Think groups, which controlled control irrelevant perceptual, memory, and semantic processes of stimuli and motor responses. We found that the believing tasks relative to thinking tasks resulted in better memory of self-related adjectives. In addition, believing compared to thinking tasks were associated with stronger activations in the left anterior insula/inferior frontal cortex and stronger functional connectivity between the mPFC and the left occipital cortex. These results provide

preliminary evidence for distinct neurocognitive processes involved in believing.

Gao et al. (2022) further tested possible cultural differences in neurocognitive processes underlying believing by collecting behavioral and fMRI responses from a Chinese and a Danish sample in the believing task used in Han et al. (2017). Reaction times and response types (yes or no responses) during believing judgments on the self and a celebrity were collected from the two cultural groups and subject to the drift-diffusion model (DDM) analyses. The results revealed three differences in cognitive processes that characterize the believing task between Chinese and Danes. First, positive and negative trait adjectives shifted the posterior distributions of the drift rate in DDM (as an index of the speed of information acquisition) either lower or larger than zero during both self- and celebrity-believing in Chinese but not in Danes. These results suggest that information acquisition during believing tasks was more sensitive to emotional contexts produced by semantic meanings of trait adjectives in Chinese. Second, the analyses of the non-decision time in DDM (as an index of processes irrelevant to decision-making) showed evidence for overlapping non-decision processes involved in self- and celebrity-believing judgments in Chinese but not in Danes. Third, the analyses of the threshold separation in DDM (as an index of decision-making strategy) suggest that the Chinese were more cautious during celebrity- than self-believing judgments whereas Danes were more cautious during self- than celebrity-believing judgments. These behavioral results are consistent with previous findings that uncovered context-dependent processing in East Asians but context-independent processing in Westerners in multiple levels of cognitive processes (e.g., Kuühnen and Oyserman, 2002; Han et al., 2011) and overlapping processes of the self and significant others in East Asians but not in Westerners (Markus and Kitayama, 1991; Zhu et al., 2007).

fMRI results of Gao et al. (2022) also revealed evidence for distinct neural activities involved in believing in the two cultural groups. Believing judgments activated the mPFC in both Chinese and Danes. However, believing judgments elicited stronger activities in the left anterior insular and ventral frontal activations in Chinese compared to Danes. In addition, Chinese participants with greater mPFC activity showed a longer duration of non-decision processes during believing-judgments. By contrast, greater mPFC activity predicted a lower degree of adopting a conservative strategy during believing judgments in Danes.

Together, the findings of cultural group differences in behavioral and neural responses during the believing task suggest that believing may be decomposed into separate processes such as information acquisition, non-decision processes, and response strategies (e.g., degree of cautiousness) that, respectively, undergo influences of individuals' cultural experiences. In addition, our results may be interpreted as that the believing task may engage deeper processing of semantic and social knowledge about others in the left ventral frontal cortex in Chinese than in Danes. Even the same brain region (e.g., mPFC) that was observed to be activated during believing in both cultural groups may be linked to different processes of believing (e.g., durations or strategies of the non-decision processes).

Discussion

The brain imaging findings summarized above have several implications for our understanding of beliefs and believing. First, people from different societies have different beliefs (e.g., different religious beliefs or different beliefs regarding the self). This cultural difference, from a neuroscience perspective, implicates that neural representations of imaginative and emotional content link an individual's prior experience with his/her future behavior vary across cultures, having in mind that this is critical for cross-cultural communications and social interactions. Second, even though there are correspondent words of "belief" and "believe" in different languages, an apparently same belief (e.g., the belief regarding the self) may have different meanings for people in different cultural environments. This implies that cognitive and neural representations of a belief may be discrepant in people from different cultures. Third, neural processes engaged in believing may also vary greatly in people from different cultures, which may reflect the consequence of social learning of beliefs that have distinct social motivations and goals in different societies.

Finally, cultural differences in beliefs are not a purely mental phenomenon but have biological underpinnings. Development of culturally specific beliefs and belief processes may be understood from the perspective of the culture–behavior–brain loop model of human development (Han and Ma, 2015; Han, 2017), which suggests both indirect culture–brain interactions, through the practice of behaviors, and direct culture–brain interactions, which constitute an interacting loop that provides a basis of human development. Shared beliefs provide a bridge to link social behavior and the brain and guide their interactions in a specific socio-cultural environment, which in turn results in the development of distinct neurocognitive processes underlying belief. Finally, it should be acknowledged that the current studies of beliefs and beliefs focused on a specific topic (e.g., self) and tested small samples of limited cultural groups. Future research may expand this line of research to other beliefs by designing new believing tasks and

collecting behavioral and brain imaging data from other large cultural samples.

Author contributions

The author confirms being the sole contributor to this work and has approved it for publication.

Funding

This study was funded by the Ministry of Science and Technology of China (2019YFA0707103), the National Natural Science Foundation of China (projects 31871134), and Chinesisch-Deutsches Zentrum für Wissenschaftsförderung (M-0093).

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Allen, R. E., Fowler, H. W., and Fowler, F. G. (1990). *The Concise Oxford Dictionary of Current English*. Oxford: Clarendon Press.

Angel, H., and Seitz, R. J. (2016). Process of believing as fundamental brain function: the concept of credition. *Forsch. Res. Bull.* 4, 1–20. doi: 10.15135/2016.4.1.1-20

Gao, T., Han, X., Bang, D., and Han, S. (2022). Cultural differences in neurocognitive mechanisms underlying believing. *NeuroImage* 250:118954. doi: 10.1016/j.neuroimage.22.118954

Grèzes, J., Frith, C. D., and Passingham, R. E. (2004). Inferring false beliefs from the actions of oneself and others: An fMRI study. *NeuroImage* 21, 744–750. doi: 10.1016/S1053-8119(03)00665-7

Klein, S. B., Loftus, J., and Burton, H. A. (1989). Two self-reference effects: The importance of distinguishing between self-descriptiveness judgments and autobiographical retrieval in self-referent encoding. *J. Pers. Soc. Psychol.* 56, 853–865. doi: 10.1037/0022-3514.56.6.853

Han, S., Mao, L., Gu, X., Zhu, Y., Ge, J., and Ma, Y. (2008). Neural consequences of religious belief on self-referential processing. *Soc. Neurosci.* 3, 1–15. doi: 10.1080/17470910701469681

Han, S., Gu, X., Mao, L., Ge, J., Wang, G., and Ma, Y. (2010). Neural substrates of self-referential processing in Chinese Buddhists. *Soc. Cogn. Affect. Neurosci.* 5, 332–339. doi: 10.1093/scan/nsp027

Han, S., and Ma, Y. (2015). A culture–behavior–brain loop model of human development. *Trends Cogn. Sci.* 19, 666–676. doi:10.1016/j.tics.2015.08.010

Han, S. (2017). *The Sociocultural Brain: A Cultural Neuroscience Approach to Human Nature*. Oxford: Oxford University Press.

Han, S., Mao, L., Qin, J., Friederici, A. D., and Ge, J. (2011). Functional roles and cultural modu- lations of the medial prefrontal and parietal activity associated with causal attribution. *Neuropsychologia* 49, 83–91. doi: 10.1016/j.neuropsychologia.2010.11.003

Han, X., Zhang, T., Wang, S., and Han, S. (2017). Neural correlates of believing. *NeuroImage* 156, 155–165. doi: 10.1016/j.neuroimage.2017.05.035

Heatherton, T. F., Wyland, C. L., Macrae, C. N., Demos, K. E., Denney, B. T., and Kelley, W. M. (2006). Medial prefrontal activity differentiates self from close others. *Soc. Cogn. Affect. Neurosci.* 1, 18–25. doi: 10.1093/scan/nsl001

Kuühnen, U., and Oyserman, D. (2002). Thinking about the self influences thinking in general: cognitive consequences of salient self-concept. *J. Exp. Soc. Psychol.* 38, 492–499. doi: 10.1016/S0022-1031(02)00011-2

Ma, Y., Bang, D., Wang, C., Allen, M., Frith, C., Roepstorff, A., et al. (2014). Sociocultural patterning of neural activity during self-reflection. *Soc. Cogn. Affect. Neurosci.* 9, 73–80. doi: 10.1093/scan/nss103

Markus, H. R., and Kitayama, S. (1991). Culture and the self: implications for cognition, emotion, and motivation. *Psychol. Rev.* 98, 224–253. doi: 10.1037/0033-295X.98.2.224

Saxe, R., and Kanwisher, N. (2003). People thinking about thinking people: fMRI investigations of theory of mind. *NeuroImage* 19, 1835–1842. doi: 10.1016/S1053-8119(03)00230-1

Seigel, J. (2005). *The idea of the self: Thought ad experience in Western Europe since the seventeenth century*. Cambridge, MA: Cambridge University Press.

Seitz, R. J. (2022). Believing and Beliefs—Neurophysiological Underpinnings. *Front. Behav. Neurosci.* 2022:16. doi: 10.3389/fnbeh.2022.880504

Sugiura, M., Seitz, R. J., and Angel, H.-F. (2015). Models and neural bases of the believing process. *J. Behav. Brain Sci.* 5, 12–23. doi: 10.4236/jbbs.2015.51002

Zhang, S. Y. (2005). *An introduction to philosophy*. Beijing: Peking University Press.

Zhu, Y., and Zhang, L. (2002). An experimental study on the self-reference effect. *Sci. China* 45, 120–128. doi: 10.1360/02yc9014

Zhu, Y., Zhang, L., Fan, J., and Han, S. (2007). Neural basis of cultural influence on self-representation. *NeuroImage* 34, 1310–1316. doi: 10.1016/j.neuroimage.2006.08.047

frontiers | Frontiers in Psychology

OPINION
published: 29 June 2022
doi: 10.3389/fpsyg.2022.943376

Learning About Credition: Exploring the Barriers Between Basic and Applied Research

*Davor Madzarevic**

Institute of Catechetics and Religious Education, University of Graz, Graz, Austria

Keywords: learning about credition, learning barriers, language use, priors, instruction about credition

OPEN ACCESS

Edited by:
Rüdiger J. Seitz,
Heinrich Heine University of Düsseldorf, Germany

Reviewed by:
Lluis Oviedo,
Pontifical University Antonianum, Italy

***Correspondence:**
Davor Madzarevic
davor.madzarevic@edu.uni-graz.at

Specialty section:
This article was submitted to Theoretical and Philosophical Psychology, a section of the journal Frontiers in Psychology

Received: *13 May 2022*
Accepted: *06 June 2022*
Published: *29 June 2022*

Citation:
Madzarevic D (2022) Learning About Credition: Exploring the Barriers Between Basic and Applied Research. Front. Psychol. 13:943376. doi: 10.3389/fpsyg.2022.943376

INTRODUCTION

This paper offers the first insights into a research project in progress: "Blind Spot Credition: Bridging the gap between basic research and application" at the Karl-Franzens University of Graz. The intended project is part of the Credition Research Project (Angel, 2022) and has a special focus on religious education, but we understand our research as paradigmatic and applicable for any kind of education dealing with the topic of belief and believing in public schools. The project has in mind the situation of teachers in public schools and has the intention to deal with learning barriers in the first approach to credition. An important aim of the project is to detect, define and analyze barriers that prevent "newcomers" (in this case teachers or school children, who haven't heard or learned about creditions so far) when encountering creditions for the first time—be it as an idea, as a concept, as a model or in theoretical debates. We argue that addressing this issue helps to bridge the gap between basic and applied research of the credition research project.

For the empirical part, we have chosen teachers of religious education in public schools as our target group. We intend to start a survey that is international (schools from Croatia, Germany, and Austria are involved in the cooperation network) and will nevertheless provide comparable results. A broader empirical database is not yet available, but we have data from pre-test studies that have not yet been published.

In this paper, I will present the theoretical background and first results which have been influencing the actual research perspectives.

LEARNING AND LEARNING-THEORIES

Learning and learning theories are fundamental when it comes to any matter that needs to be mastered hermeneutically. We do not intend to contribute to the theory-building of learning theories. When we talk about learning, we mean a strategy on how to overcome super complex material burdened with various hermeneutic barriers. Our basic assumption is that without detecting and deciphering those barriers our learning approach to credition will be finished before it effectively starts. We are interested in the emotional foundations of barriers and their learning implications.

We also want to understand which ways or means are the most appropriate to help school children and their teachers to understand the basic concept of credition. We want to figure out how pathways can be developed to make credition attractive to them, even though they do not and cannot know what to expect.

Finally, in an attempt to identify the barriers, we are interested in the research of timeline: Starting condition, learning steps, their challenges, and the end condition.

DIFFERENT STARTING CONDITIONS

Starting condition refers to the three different levels from which our research proceeds: school children, teachers, and science.

Based on pre-test experiences, school children most often find the material from the credition research project uncomfortable because they automatically associate it with religion, but also with prejudices related to, e.g., philosophy. They often complain about the problem of hermeneutics and the complexity of the credition topic as well. Teachers also encounter a variety of issues in learning about credition. They mostly find a tense relationship between motivation, energy investment in learning, and the benefits of credition. Learning about credition does not progress smoothly in the scholarly fields either. Those who come from the Humanities often find it daunting that there is a wide range of interdisciplinary knowledge about credition. Different disciplines are included: philosophy, cognitive science, and natural science. And sometimes it can be difficult for people to understand that there is conceptual knowledge as the basis for understanding the model.

There are also a lot of incompatibilities and cross relations between different starting positions. For example, when speaking about beliefs, school children or teachers may have in mind religion but not epistemology. Epistemologists may be thinking about philosophical issues, i.e., justification but not the need to make these debates accessible, etc.

One of the first empirical pre-test results is that there are some difficulties and complications in the understanding of credition, and therefore finding a starting point for the research represents a crucial challenge. We assume that the starting point cannot be credition itself, but human consciousness. Exploring associations of "belief" from pre-test studies seemed fruitful to our research reflections. In the pre-tests, we found the word association method (Kent and Rosanoff, 1910) helpful. In our research approach, we will highlight parameters that indicate credition as a blind spot for learning. One of these is how language is used when talking about beliefs.

LANGUAGE USE

Pre-test studies show that language use contributes to barriers in first learning about credition. Thus, in a pre-test study conducted in a primary school in Bosnia and Herzegovina, with the sample of 27 pupils, mostly the following associations to the term "belief" popped up: God (came up even 15 times), Jesus, love, hope, holiness, prayer, etc. When asked what "belief" means to them, the students mostly answered: "Belief means to have belief in someone, for example, to trust your parents or to believe in God" or "It means a lifestyle in which we are devoted to God". In two other similar questionnaires more than half of the responses indicated linking believing with religion. These associations and definitions of belief have certain implications for at least the three main characteristics of understanding the belief:

1. Belief as a noun.
2. Belief as religious belief.
3. Belief concerning (religious) content.

This also suggests that learning about credition has to do with the students' different backgrounds in learning subjects, their predispositions to learn, their worldviews, and especially their attitudes toward "beliefs". Specifically, this means that how the language is used indicates cognitive assumptions of prior knowledge and information that can influence the adoption of new information, knowledge, and new conclusions, which in psychology are referred to as "priors" (Tobias, 1994; Dochy, 1996).

PRIORS

In our research we have already identified some consequences of different priors related to the above-mentioned common use of beliefs:

(1) In a broader philosophical sense, the question of belief is embedded in a long tradition of Western thinking and has produced a rich and overwhelmingly broad literature base over about 2,500 years. Therefore, one may get the impression that belief is a well-defined phenomenon, but newer interdisciplinary approaches to the processes of believing deny such an understanding and show that belief is an ill-defined phenomenon.

(2) Another prior may be identified in an approach to belief which seems to be especially influential in neuropsychiatry and psychology. Belief seems to be associated with pathology. Thus, belief can be related to neurosis or delusion (McCauley and Graham, 2020). This can cause a variety of problems because linking pathology to belief can automatically cause a negative attitude toward any approach to credition.

(3) The everyday use of language, also demonstrated in pre-test studies, shows a close connection between understanding "belief" as a religious belief. Evidence for this can be also found on the theoretical level in the credition literature: "No other concept relevant to understanding human behavior is as deeply tied to religion as belief" (Angel et al., 2017, p. 5). This is of course highly problematic for a correct understanding of credition that is not limited just to religion.

(4) Another prior is the very frequent use of "belief" both in everyday speech and in scientific discourse as a noun. For instance, the predominant use of nouns like "formation of belief" (Langdon and Connaughton, 2013) or "dynamics of belief" (Forrest, 1986) work against conceptualizing believing processes as having a fluid character (cf. Angel, 2017, p. 19).

(5) In everyday understanding, beliefs are often content-oriented. Testimonial beliefs (I believe in) and fiduciary beliefs (to have faith in) are often expressed here. Such a prior significantly reduces the likelihood of a proper understanding of the process of believing.

DISCUSSION

Barriers to an Approach to Credition

The core task of the project in progress, which this paper intends to present, is to illuminate barriers that prevent a correct

understanding of the fluidity of believing in the school context. To achieve the desired goal, we use already existing scientific parameters such as:

epistemology, which focuses predominantly on the question of justification (Runehov, 2017); philosophy of mind, which focuses predominantly on the nature of beliefs (Visala and Angel, 2017) and eliminativism, which claims that belief should not be a matter of scientific debate (Stich, 1996), etc.

These parameters influence the several theoretical levels that coincided with some results which were obtained from pre-test studies:

(1) Process of believing is a blind spot in the mind; therefore, no initial associations are pointing to the fluidity of beliefs.
(2) Belief is initially marked as a noun in the mind. Therefore, the paradigmatic shift from belief to credition seems to be irrelevant.
(3) Credition has something to do with belief. Belief is immediately associated with religion, therefore credition is monopolized by religion. The same confusion is with the term "religiosity" (cf. Angel in this volume) which seems instinctively associated with religion. This is counterproductive for any understanding of the processes of believing. From a cognitive neuroscience point of view, it must be stated that creditions do not take place in religions but rather in humans when they develop and live their religiosity.
(4) Belief has something to do with knowledge, but the relevance of epistemological discussions seems to play a significant role for the newcomers.
(5) The initial notion of cognitive science and neuroscience is the neglect of the mind. Even if belief is understood as an inner process, then biological and cognitive science background knowledge is required and already provided by science.

Barriers to the Instruction About Credition

Addressing barriers in this project has several implications for a particular strategy for instruction about credition. In our approach, we trace the following strategies. First, it is necessary to draw attention to credition as a blind spot. Second, it is necessary to make the blind spot attractive enough to provoke energy and exertion for learning. Finally, it is important to make attractive the benefit of learning about credition for pupils. A special emphasis could be placed on those aspects where the topic of belief comes directly or indirectly into play, such as the role of creditions in dealing with catastrophes (Sugiura, 2017) or the connection between creditions and identity development (Colagè and Gobbi, 2017) or the influence of creditions on decision-making (Hick et al., 2020), etc.

Following these three important pre-steps, it is possible to anticipate further strategic steps for learning. Explaining the scientific background of the credition concept (cf. Angel in this volume) can represent one of the initial steps. In developing different strategic steps, setting clear goals for learning about credition should not be neglected as well. Determining the amount of information and knowledge within a certain time frame to achieve the desired goals is therefore of primary interest in the learning strategy. At the same time, the balance between investing energy and achieving set goals should be kept in mind.

Identifying and analyzing barriers in the approach to credition enables didactic creativity in presenting the concept and model to pupils as well. In doing so, teachers should pay attention to avoiding already established barriers and provide students with the most unobtrusive approach to credition. They could also use various didactic methods and means of digital learning and teaching when presenting the concept and the model. This would make the matter they are learning as interesting as possible. For that aim, E-Learning Methodologies and Tools (Wang, 2012) based on Cognitive Load Theory (Sweller et al., 2019) could be useful here. Finally, teachers should motivate and enable students to work individually with the model of credition and encourage communication of their personal experiences and reflections in working with the model (cf. Mitropoulou et al., 2018).

In the end, developing a learning and teaching strategy should help to integrate credition more successfully into the future school context. Some of the theoretical steps presented here are part of this paper, but we can only evaluate the results of the study after they are available.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

This article was funded by Dr. Rüdiger Seitz, via the Volkswagen Foundation, Siemens Healthineers, and the Betz Foundation.

ACKNOWLEDGMENTS

The author would like to express my very great appreciation to Prof. Dr. Hans-Ferdinand Angel for his valuable and constructive suggestions during the planning and development of this research work.

REFERENCES

Angel, H. F. (2017). "Credition: from the question of belief to the question of believing," in *Processes of Believing: The Acquisition, Maintenance, and Change in Creditions*, eds H. F. Angel, L. Oviedo, R. F. Paloutzian, A. L. Runehov, and R.J. Seitz (Cham: Springer), 17–36.

Angel, H. F. (2022). *The Credition Research Project is an Interdisciplinary and International Project That Contains a Comprehensive Theory of Believing Processes.* Available online at: https://credition.uni-graz.at/de/ (accessed June 04, 2022).

Angel, H. F., Oviedo, L., Paloutzian, R. F., Runehov, A. L., and Seitz, R. J. (eds.). (2017). "Introduction: what we do not know about believing – approaching a new scientific hot spot," in *Processes of Believing: The*

Acquisition, Maintenance, and Change in Creditions (Cham: Springer), 3–15.

Colagè, I., and Gobbi, N. (2017): "Creditions and self-identity: embodiment between actions and (life) decisions," in *Processes of Believing: The Acquisition, Maintenance, and Change in Creditions*, eds H. F. Angel, L. Oviedo, R. F. Paloutzian, A. L. Runehov, and R. J. Seitz (Cham: Springer), 247–262.

Dochy, F. J. R. C. (1996). "Assessment of domain-specific and domain-transcending prior knowledge: entry assessment and the use of profile analysis," in *Alternatives in Assessment of Achievements, Learning Processes and Prior Knowledge*, eds M. Birenbaum, and F. J. R. C. Dochy (Boston, MA: Kluwer), 227–64.

Forrest, P. (1986). *The Dynamics of Belief*. Oxford: Blackwell.

Hick, H., Angel, H.-F., Kranabitl, P., and Wagner-Skacel, J. (2020): "Decision-making and the influence of the human factor," in *Systems Engineering for Automotive Powertrain Development*, eds H. Hick, K. Küpper, and H. Sorger (Cham: Springer), 1–26.

Kent, G. H., and Rosanoff, A. J. (1910). A study of association in insanity. *Am. J. Insanity* 67, 37–96, 317–90. doi: 10.1037/13767-000

Langdon, R., and Connaughton, E. (2013). "The neuropsychology of belief formation," in *The Neural Basis of Human Belief Systems*, eds F. Krüger, and J. Grafman (Hove: Psychology Press), 19–42.

McCauley, R. N., and Graham, G. (2020). *Hearing Voices and Other Matters of the Mind: What Mental Abnormalities Can Teach Us About Religions*. Oxford: Oxford University Press.

Mitropoulou, V., Gkirlou, D., and Meke, M. (2018). Building worldviews with credition teaching model in high school curricula of religious education: a controlled experimental study. *Soc. Study* 8, 325–335. doi: 10.17265/2159-5526/2018.07.003

Runehov, A. L. C. (2017). "Credition and justification," in *Processes of Believing: The Acquisition, Maintenance, and Change in Creditions*, eds H. F. Angel, L. Oviedo, R. F. Paloutzian, A. L. Runehov, and R. J. Seitz (Cham: Springer), 195–207.

Stich, S. (1996). *From Folk Psychology to Cognitive Science: The Case Against Belief*. Cambridge: MIT Press.

Sugiura, M. (2017). "The power to live with disasters: adaptive believing processes of the self and world," in *Processes of Believing: The Acquisition, Maintenance, and Change in Creditions*, eds H. F. Angel, L. Oviedo, R. F. Paloutzian, A. L. Runehov, and R. J. Seitz (Cham: Springer), 139–152.

Sweller, J., Van Merriënboer, J. J. G., and Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educ. Psychol. Rev.* 31, 261–292. doi: 10.1007/s10648-019-09465-5

Tobias, S. (1994). Interest, prior knowledge, and learning. *Rev. Educ. Res.* 64, 37–54. doi: 10.3102/00346543064001037

Visala, A., and Angel, H. F. (2017). "The theory of credition and philosophical accounts of belief: looking for common ground," in *Processes of Believing: The Acquisition, Maintenance, and Change in Creditions*, eds H. F. Angel, L. Oviedo, R. F. Paloutzian, A. L. Runehov, and R. J. Seitz (Cham: Springer), 183–193.

Wang, V. C. (2012). Understanding and promoting learning theories. *Int. J. Multidiscipl. Res. Mod. Educ.* 8, 343–347.

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Madzarevic. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Psychology

TYPE Perspective
PUBLISHED 21 September 2022
DOI 10.3389/fpsyg.2022.926489

Check for updates

OPEN ACCESS

EDITED BY
Hans Ferdinand Angel,
University of Graz, Austria

REVIEWED BY
Lluis Oviedo,
Pontifical University Antonianum, Italy

*CORRESPONDENCE
Oliver Davies
oliver.davies@kcl.ac.uk

SPECIALTY SECTION
This article was submitted to Theoretical and Philosophical Psychology, a section of the journal Frontiers in Psychology

RECEIVED 22 April 2022
ACCEPTED 25 July 2022
PUBLISHED 21 September 2022

CITATION
Davies O (2022) The architecture of creditions: Openness and otherness.
Front. Psychol. 13:926489.
doi: 10.3389/fpsyg.2022.926489

COPYRIGHT
© 2022 Davies. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

The architecture of creditions: Openness and otherness

Oliver Davies[1,2,3]*

[1]King's College London, London, United Kingdom, [2]LSRI Research Institute at Campion Hall, University of Oxford, Oxford, United Kingdom, [3]Wolfson College, University of Oxford, Oxford, United Kingdom

"Creditions" are an important new idea within our contemporary understanding of the human. They potentially represent the unity of both humanistic and scientific ways of modeling the human. As such, "creditions" offer a bridge between current thinking in science and the humanities and the development of a more powerfully integrated interdisciplinary hermeneutic. It is argued in this article that the questions posed by "creditions" (as developed by Rüdiger Seitz and Hans-Ferdinand Angel) cannot be resolved through reduction but rather only through cohesive systematization. In contrast with coherence in conventional science, "credition-centered" thinking finds expression in systemic ways. The complex humanity of the reflective subject resists reduction; and calls to be analyzed in terms of sociality, the identification of "otherness" and interactive engagement. In this context then a thinking which is attuned to complexity and to otherness has an important place in the expression of the social subject as a complex and relational self, in today's world. These are not however social realities as we find them either in large-scale social schemata, or indeed in the intimacy of the face to face. Rather credition-centered learning falls between these two categories and is best described as "the productive knowledge of community," where community is generated by productive enhancement and the embrace of otherness over time.

KEYWORDS

community, otherness, creditions, interdisciplinarity, integration, freedom, Franciscans, belonging

Introduction: Contribution to an architecture of creditions

The concept of credition is an important new idea within our contemporary understanding of the human (Angel et al., 2017; see also Seitz and Angel, 2020). Creditions potentially represent the unity of both humanistic and scientific ways of modeling the human. As such, the concept of credition offers a bridge between current thinking in science and the humanities and the development of a more powerfully integrated interdisciplinary hermeneutic. The complex humanity of the reflective subject resists reduction; and calls to be analyzed in terms of sociality, the identification of "otherness" and interactive engagement. In this context then a thinking which is attuned to complexity and to otherness has an important

place in the expression of the social subject as a complex and relational self (Seligman et al., 2008; Zeman, 2009; Di Paolo and De Jaegher, 2012; Han, 2017). These are not however social realities as we find them either in large-scale social schemata, or indeed in the intimacy of the face-to-face. As the concept of credition embraces processes in individuals as well as in groups or societies it is argued in this article that the concept of credition encourages a cohesive systematization of human behavior. In contrast with coherence in conventional science, "credition-centered" thinking finds expression in systemic ways while credition-centered learning might be best described as "the productive knowledge of community," where community is generated by productive enhancement and the embrace of otherness over time (Gallagher, 2008; Tononi et al., 2016; Bente and Novotny, 2020).

Therefore this article intends to contribute to the "Architecture of Creditions" from a specific perspective which focusses on the combination of different poles, namely "Openness and Otherness." From this perspective the terms "Openness" and "Otherness" together are understood to be key aspects for a definition of creditions. This article seeks to address a far-reaching problem concerning the nature of productive human relations. The academy is used to large scale population studies on the one hand and to small scale (face-to-face) sociality on the other. The intervening level of extended intimacy or productive sociality is far less present. And yet this is the level at which our "belonging" appears and is stabilized. It is the domain of our integrated identity. It is here then that the concept of "creditions" has a critical role to play. "Credition-theory" allows the emergence of "otherness" as a form of social openness. This in turn opens up to the sphere of ritual, in which the material properties of the linguistic sign, as shape and sound, are celebrated, in accordance with the presence of our advanced linguistic consciousness (Bell, 1992; Konvalinka and Roepstorff, 2012; Ramstead et al., 2016). Credition theory however enables the unfolding of a further analytical stage. This is the development of our understanding of the linguistic sign as mediating freedom. Credition theory can offer the realization of a typology of freedom, as a key factor in the development of our self-understanding, through the embrace of "openness" and "otherness" (Anderson, 2016). Integrating both concepts seems to be the basic challenge for learning the higher prosocial level.

Beliefs as results of believing and believing processes

The credition concept highlights the dynamic of believing processes which result in mental representations which might be called beliefs. One of the innovative aspects of this approach results from neurophysiological findings which focus on specific believing processes (Seitz et al., 2018; Seitz and Angel, 2020). Three types can be distinguished. These are labeled as empirical, relational, and conceptual beliefs. These processes contribute in mutual interaction to the production of beliefs. My focus will be on conceptual beliefs. They are language-bound, narrativist and participative; and they involve ritual. This generates a stance of "believing in." "[G]iven the involved neural processes of meaning-making and affective loading, conceptual beliefs appear similar to empirical and relational beliefs but are far more abstract" (Seitz and Angel, 2020, p. 3).

The capacity to develop more complex believing processes can be seen as the result of brain evolution. "The neural processes underlying formation and maintenance of beliefs in an increasingly complex social environment demanded augmented processing resources in the brain" (Fuentes, 2017; Seitz and Angel, 2020, p. 3–4). There is evidence for the possibility that "this enhanced processing demand was the force driving the phylogenetic enlargement of the parietal and frontal cortex which are key cortical areas in cerebral circuits affording integrative supramodal information processing" (Seitz and Angel, 2020, p. 3–4). The crucial further point here is that human complexity points to choice and so also to the complex phenomenon of freedom.

Evolution, rituals, and tool use

Seitz and Angel propose that there is a consistent link between "conceptual beliefs" and "ritual," whereby multi-modal complexity is constantly enhanced (Whitehouse, 2021). But what is the most concrete evolutionary source of this complexity? It has been proposed that the so-called "ratcheting effect" (Tennie et al., 2009) has played a key role whereby two different orientations in the world – interfacial orientation and hand-world tool use in combination – generated a new system which itself represents enhanced creativity. Advanced linguistic consciousness then is based in the interplay of the human interface and sociality on the one hand, and tool use or technology on the other. These are powerful, rotating, evolutionary drivers. From this perspective, words can be defined as "social tools" which combine sociality and technology in their original pre-modern setting.

Recent experiments in the learning of stone tool-making techniques reinforce the role of technology in the origins of language (Hurford, 2007; Lombao et al., 2017). Clark (2011) has pointed to the ways in which language and stone use mirror each other. In turn, Jayne Wilkins has argued for the emergence of "dialects" in key areas of the Still Bay and Howieson's Port in Southern Africa, on the basis of "imitative social learning" and discrete sets of "stone tool technological traits." Wilkins argues persuasively that distinctive sequences of strikes but also of the sounds of tool-making developed, and were expressed as distinct "dialects" or "schools" which paralleled the emergence of distinctive linguistic dialects (Stout and Chaminade, 2012; Wilkins, 2020; Dunbar, 2022).

But how are we shaped by this inheritance today? Firstly, according to Saussurean linguistics from the early 20th century, each utterance (parole) requires a choice between a range of

potential linguistic possibilities (langue). We choose our words from all the available words we might have used, and so we allow ourselves to be held to account for them (De Saussure, 1986). Saussurean linguistics reinforces the role of freedom therefore, as arising from the internalization of external tools, through life-long practices of speaking and writing which together constitute our Advanced Linguistic Consciousness or "ALC" (Chalmers, 2010; Huth et al., 2016).

Re-reading historical concepts with the modern lenses of cognitive science

It is increasingly evident today that there are no grounds for uncoupling our positivist, controlling freedom "from" and freedom "to" from the strongly consensual, rhythmic social modalities of our human "social cognition," which is our freedom "in" (Schilbach et al., 2013; Bente and Novotny, 2020; Davies, 2021). Indeed, this broader integration arguably marks the point of a deeper humanization, and indeed is perhaps the locus of our power of choice. But we need to take note too of the effect of "otherness." Creditions theory allows the coexistence of a community at Time "A" and Time "B." Time "A" might be the launch of the Franciscan community in the 13th century with records of its compelling need to come to judgment about this new, enriched but also very challenging form of ethical life. Time "B" on the other hand may be the current reader's own time framework. In Time "B," those who have been influenced by contemporary community-based credition theory may well empathize with the records and data of Time "A." It may be that Time "B" and Time "A" can interact with one another, as Time "B" discerns the "otherness" of Time "A" and begins openly to engage with it and to learn from it, in the formation of a trans-historical community based upon the reception of a productive "otherness." The productivity of "creditions" needs to be grounded both in the cultural and the historical forms of our sociality, on the one hand, and in the contemporary science of human sociality, on the other.

Dante (1265–1321) offers the classical, transformational definition of language, which is that language is both sensuale and rationale. This means that "language, as a system of visible or oral signs, reproduces that peculiarly human mix that we ourselves are, of matter and mind, materiality and conceptuality," reflecting the concept of "rational animality" as developed by Thomas Aquinas (Turner, 2004, p. 89–93; Davies, 2015, p. 248). In De vulgari eloquentia Dante writes: "It is more truly human for a human being to be perceived than to perceive" (Botterill, 1996, 1.3.7). The Divine Comedy is the cosmic enactment of that reality which is, as such, deeply consistent with the integrated science of our own times. Here Dante offers us a profound image of our "freedom in" on a cosmic scale which parallels current thinking on the role of the materiality of language in human relationality and human cognition.

TABLE 1 Medieval and modern concepts of sociality.

Identifier	Original theme	Keyword	Modern application	Level of correspondence (1–5)
3d (2I)	"Thisness"	Haecceitas	Immersiveness	(4)
3d (2II)	"Natural Law"	Impressa	Social cognition	(5)
3d (2III)	"Projected sociality"	Condilectio	Prosociality	(5)
3d (2IV)	"Decision-making"	Non-velle	Symmetry	(4)

This table represents the comparison between early Franciscan notions of the social and current scientific conceptions of the social on a scale of 1–5. 1 represents minimal similarity and 5 represents extensive similarity.

The text known as the Summa Halensis (SH) was collaboratively authored by the founding members of the Franciscan school at Paris (1236–1245). It was not only the first official statement of Franciscan thought but also became a defining text which explored fundamental distinctions between philosophy and theology (Saccenti, 2020; Schumacher, 2020). It is this text, together with the later writings of the Franciscan scholar Duns Scotus, which appear to break new ground in understandings of the long-term practices of human sociality as manifest in "immersiveness" (4d i), "social cognition" (4d ii), "prosociality" (4d iii), and "symmetry" (4d iv) (see **Table 1**). (4d i–iv) can all be identified as modes of openness toward otherness. These are core representations of creditions as ways of integrating openness within complexity.

Haecceitas – "immersiveness"

Scotus roots his anthropology in space and time and in our embodied human particularity. But he also develops an innovative metaphysics of particularity or what he calls haecceitas ("this-ness"). Haecceitas signals that we cannot define real things through the language either of "matter" or "form" alone, but neither can we define them through "matter and form" in combination, as was the norm. This also is too abstract. Rather, haecceitas points to real things as being a particular combination of both "matter" and "form" in this space and time. This finds parallels among contemporary philosophers today (e.g., Dancy, 2018). Scotus' emphasis on particularity and "this-ness" yielded a new kind of metaphysics, one which participates, for Scotus, in the beauty of the original divine creation. This points to the otherness of the particular.

Impressa – "social cognition"

The SH argues decisively for an account of morality which is based in "natural law." Once again the thinking is physicalist: "natural law is knowledge of the eternal law impressed in the soul." Here our sense of morality is a given. The early Franciscans argued that "the eternal law is received by rational creatures and thus it is made present to their minds through impression rather than through an autonomous search on the part of reason itself." (Saccenti, 2020, p. 227–250). This

is arguably a physicalist ethics or an ethics of embodiment. It corresponds well with our own contemporary accounts of the role of the social cognition system as embedded, and as constituting the active ground of our social understanding and social bonding.

Condilectio – "prosociality"

The Franciscans were drawn in particular by the concept of condilectio as "shared love" or "co-love," which they also understood to be related to a "love of justice." As Lydia Schumacher states: "co-love occurs when a third is loved by the two in harmony and collectively (concorditer et socialiter) so that the two persons' affects are fused to become one because of the flame of love for the third." (Schumacher, 2019, p. 174). In its original context this is a version of Trinitarian theology, which places a particular emphasis on the "third" beyond the dyad of the "inter-face." But we can also read this today as proposing "love" as a form of radical openness which is actualized beyond the "interface." This appeals to the extension of love, as based in the social cognition system, into larger scale society, along the axis of a universalist "love for justice."

Non-velle – "symmetry"

The Franciscan vocation itself (which involved a vow of radical poverty) focused the minds of leading intellectuals, and especially Duns Scotus, on the nature of decision-making. For Scotus, three kinds of freedom predominate: velle ("I want"), nolle ("I don't want") and non-velle ("my mind is still open") (Ingham and Dreyer, 2004, p. 146–172). Velle and nolle both point to a form of self-interested possessiveness (affectio commodi), while the third points to our preparedness to remain detached and open in our moral questioning. Scotus calls this affectio iustitiae, or "love for justice."

Here parallels emerge between openness in decision-making as Scotus and the early Franciscans develop it (Schumacher, 2020) and the neurological work, for instance, of Robert Kane. Kane describes how ethical challenges are represented in "movement away from thermal equilibrium – in short a kind of stirring up of chaos in the brain that makes it sensitive to micro-indeterminacies at the neuronal level." Kane observes that the brain is a kind of "parallel processor (. . .) which can simultaneously process different kinds of information relevant to tasks such as perception or recognition through different neural pathways." This processing capacity is "essential to the exercise of free will." Kane adds that "[t]he key to difficult ethical decision-making, in which none of the initial possibilities appear to allow resolution, is time, effort and finally the formation of new neural pathways in the brain through the top-down effect." These create the possibility of a new future and identity, and they constitute "growth" (Kane, 2011).

Discussion

We can postulate that "creditions" can ultimately be defined in terms of the openness of the self as emergent within evolutionary contexts, involving "the phylogenetic enlargement of the parietal and frontal cortex which are key cortical areas in cerebral circuits affording integrative supramodal information processing" (Seitz and Angel, 2020, p. 3–4). The credition-centered thematization of complexity which is undertaken in the present project itself constitutes an openness to, or within, complexity. Furthermore, this openness bears the characteristics of freedom, or play, as an originary and fundamental characteristic of the human. Playing together is one of the key ways in which we develop and express our humanity. In particular, play can also be characterized in terms of freedom, or irrepressible non-compulsion.

It is this aspect of "freedom," within an "architecture of complexity and otherness," which begins to open up the possibility that the hermeneutics developed within creditions-theory may also overlap more directly with other forms of human self-possession; and specifically with that kind of self-possession which we can identify with the self's belonging. Our social belonging is grounded in our acceptance by the other. The free movement of the other is prior. But this points to a further configuration, which is the foundational role of co-ordinated movement within relationality, as we speak and interact with each other, not least through maintaining eye contact. If they are viewed from another perspective and from within a different set of presuppositions, the spontaneity of such movements can be judged from the perspective of the terminology of ritual and repetition. Here it may appear that the freedom of movement we associate with the spontaneity of formal ritual can re-emerge as a form of life, and so contribute to a new phase in our human self-understanding.

Author contributions

The author confirms being the sole contributor of this work and has approved it for publication.

Funding

This manuscript was funded by Rüdiger Seitz, via the Volkswagen Foundation, Siemens Healthineers, and the Betz Foundation. Siemens Healthineers was not involved in the study design, collection, analysis, interpretation of data, the writing of this article, or the decision to submit it for publication.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Anderson, E. (2016). "Freedom and equality," in *The Oxford Handbook of Freedom*, eds D. Schmidtz and C. Pavel (Oxford: Oxford University Press). doi: 10.1093/oxfordhb/9780199989423.013.25

Angel, H.-F., Oviedo, L., Paloutzian, R. F., Runehov, A. L. C., and Seitz, R. S. (2017). *Processes of Believing: The Acquisition, Maintenance, and Change in Creditions*. Berlin: Springer Verlag. doi: 10.1007/978-3-319-50924-2

Bell, C. (1992). *Ritual Theory, Ritual Practice*. Oxford: OUP.

Bente, G., and Novotny, E. (2020). "Bodies and minds in sync: forms and functions of interpersonal synchrony in human interaction," in *Handbook of Communication, Science and Biology*, eds K. Floyd and R. Weber (New York, NY: Routledge), 416–428. doi: 10.4324/9781351235587-36

Botterill, S. (1996). *Dante: De Vulgarii Eloquentia*. Cambridge, MA: Cambridge University Press.

Chalmers, D. (2010). *The Character of Consciousness*. Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780195311105.001.0001

Clark, A. (2011). *Supersizing the Mind: Embodiment, Action and Cognitive Extension*. New York, NY: Oxford University Press. doi: 10.1007/s11098-010-9598-9

Dancy, J. (2018). *Practical Shape: A Theory of Practical Reasoning*. Oxford: Oxford University Press. doi: 10.1093/oso/9780198805441.001.0001

Davies, O. (2015). ""Love was his meaning": on learning from medieval texts," in *Desire, Faith, and the Darkness of God*, eds E. Bugyis and D. Newheiser (Notre Dame: University of Notre Dame Press), 242–267. doi: 10.2307/j.ctvpj7dcs.15

Davies, O. (2021). "Science, philosophy and the authority of the early franciscan summa halensis: learning from the past for the sake of the future," in *The Legacy of Early Franciscan Thought*, ed. L. Schumacher (Berlin: De Gruyter), 373–397. doi: 10.1515/9783110684827-021

De Saussure, F. (1986). *Course in General Linguistics*, 3rd Edn. Chicago, IL: Open Court.

Di Paolo, E., and De Jaegher, H. (2012). The interactive brain hypothesis. *Front. Hum. Neurosci.* 6:21.

Dunbar, R. (2022). *How Religion Evolved and Why it Endures*. London: Pelican Books.

Fuentes, A. (2017). *The Creative Spark: How Imagination Made Humans Exceptional*. Dutton, MT: Penguin Publishing Group.

Gallagher, S. (2008). "Understanding others: embodied social cognition," in *Handbook of Cognitive Science: An Embodied Approach*, eds P. Calvo and T. Gomila (San Diego, CA: Elsevier), 439–452.

Han, S. (2017). *The Sociocultural Brain: A Cultural Neuroscience Approach to Human Nature*. Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780198743194.001.0001

Hurford, J. (2007). *The Origins of Meaning. Language in the Light of Evolution*. Oxford: Oxford University Press.

Huth, A. G., de Heer, W. A., Griffiths, T. L., Theunissen, F. E., and Gallant, J. L. (2016). Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature* 532, 453–458. doi: 10.1038/nature17637

Ingham, M. B., and Dreyer, M. (2004). *The Philosophical Vision of John Duns Scotus*. Washington, DC: Catholic University of America Press.

Kane, R. (2011). "Rethinking free will: new perspectives on an ancient problem," in *The Oxford Handbook of Free Will*, 2nd Edn, ed. R. Kane (Oxford: Oxford University Press). doi: 10.1093/oxfordhb/9780195399691.003.0019

Konvalinka, I., and Roepstorff, A. (2012). The two-brain approach: how can mutually interacting brains teach us something about social interaction? *Front. Hum. Neurosci.* 6:215. doi: 10.3389/fnhum.2012.00215

Lombao, D., Guardiola, M., and Mosquera, M. (2017). Teaching to make stone tools: new experimental evidence supporting a technological hypothesis for the origins of language. *Sci. Rep.* 7:14394. doi: 10.1038/s41598-017-14322-y

Ramstead, M., Veissière, S., and Kirmayer, L. (2016). Cultural affordances: scaffolding local worlds through shared intentionality and regimes of attention. *Front. Psychol.* 7:1090. doi: 10.3389/fpsyg.2016.01090

Saccenti, R. (2020). "From the "lex aeterna" to the "leges addictae"," in *John of La Rochelle and the Summa Halensis,' in The Summa Halensis: Doctrines and Debates*, ed. L. Schumacher (Berlin: De Gruyter), 227–250. doi: 10.1515/9783110685008-015

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., et al. (2013). Toward a second-person neuroscience. *Behav. Brain Sci.* 36, 393–414. doi: 10.1017/S0140525X12000660

Schumacher, L. (2019). *Early Franciscan Theology: Between Authority and Innovation*. Cambridge, MA: Cambridge University Press.

Schumacher, L. (ed.) (2020). *The Summa Halensis: Doctrines and Debates: Veröffentlichungen des Grabmann-Institutes zur Erforschung der mittelalterlichen Theologie und Philosophie*. Berlin: De Gruyter.

Seitz, R. J., and Angel, H.-F. (2020). Belief formation – A driving force for brain evolution. *Brain Cogn.* 140:105548. doi: 10.1016/j.bandc.2020.105548

Seitz, R. J., Paloutzian, R., and Hans-Ferdinand, A. (2018). From believing to belief: a general theoretical model. *J. Cogn. Neurosci.* 30, 1254–1264.

Seligman, A. B., Weller, R. P., Puett, M., and Simon, B. (2008). *Ritual and Its Consequences: An Essay on the Limits of Sincerity*. Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780195336009.001.0001

Stout, D., and Chaminade, T. (2012). Stone tools, language and the brain in human evolution. *Philos. Trans. R. Soc. B Biol. Sci.* 367, 75–87.

Tennie, C., Call, J., and Tomasello, M. (2009). Ratcheting up the ratchet: on the evolution of cumulative culture. *Philos. Trans. R. Soc. B Biol. Sci.* 364, 2405–2415. doi: 10.1098/rstb.2009.0052

Tononi, G., Boly, M., Massimini, M., and Koch, C. (2016). Integrated information theory: from consciousness to its physical substrate. *Nat. Rev. Neurosci.* 17, 450–461. doi: 10.1038/nrn.2016.44

Turner, D. (2004). *Faith, Reason, and the Existence of God*. Cambridge, MA: Cambridge University Press. doi: 10.1017/CBO9780511617317

Whitehouse, H. (2021). *The Ritual Animal: Imitation and Cohesion in the Evolution of Social Complexity*. Oxford: Oxford University Press. doi: 10.1093/oso/9780199646364.001.0001

Wilkins, J. (2020). "Archaeological evidence for human social learning and sociality in the middle stone age of South Africa," in *Theology and Evolutionary Anthropology. Dialogues in Wisdom, Humility and Grace*, eds C. Deane-Drummond and A. Fuentes (London: Routledge), 119–141. doi: 10.4324/9780429273650-10

Zeman, A. (2009). *A Portrait of the Brain*. London: Yale University Press.

frontiers in Psychology

MINI REVIEW
published: 21 March 2022
doi: 10.3389/fpsyg.2022.817940

Categorical Versus Graded Beliefs

*Franz Dietrich**

Paris School of Economics, CNRS, Paris, France

This essay discusses the difficulty to reconcile two paradigms about beliefs: the binary or categorical paradigm of yes/no beliefs and the probabilistic paradigm of degrees of belief. The possibility for someone to hold beliefs of both types simultaneously is challenged by the lottery paradox, and more recently by a general impossibility theorem by Dietrich and List. The nature, relevance, and implications of the tension are explained and assessed.

Keywords: logic vs. rational choice theory, yes/no belief vs. graded belief, subjective probability, belief binarization, lottery paradox, impossibility theorem, binary belief, credence

1. TWO TYPES OF BELIEF AND THEIR POTENTIAL COEXISTENCE

Rational choice theory and logic have very different concepts of belief, each of which enjoys significant appeal and wide applications. Rational choice theory takes agents to have graded beliefs of the form of subjective probability assignments. One might believe that it rains with subjective probability 2/3, or that one will stay healthy with subjective probability 3/4. By contrast, logic takes agents to have categorical beliefs, of the form of "yes" or "no" (or abstention). One might believe that it rains, or that one will stay healthy, in a categorical rather than graded sense. Believing something categorically should not be confused with complete certainty, i.e., with maximal graded belief: otherwise one would hardly ever believe anything in the categorical sense.

The advantage for rational choice theory of assuming probabilistic beliefs is considerable: it opens to the door to the classic notion of a rational agent seeking to maximise expected utilities, since expected utilities are the result of combining the probabilistic model of beliefs with the utility-based model of goals, values, or desires. As such, probabilistic beliefs form an intrinsic part of the classic homo oeconomicus. By contrast, logicians are less interested in decision making, and, hence, do not need to combine beliefs with goals, values, or desires. Instead, they often focus on beliefs alone, which they usually take to be truth-oriented, logically consistent, and deductively closed, and to evolve *via* reasoning and belief revision. Categorical beliefs lend themselves to reasoning and belief revision, as logicians have amply demonstrated.

Of course, rational choice theory has its own theory of belief revision: a highly unified Bayesian theory, in which probabilistic beliefs undergo Bayesian updating as new information arrives. But it is questionable whether Bayesianism yields a theory of reasoning as opposed to revision, and more generally whether probabilistic beliefs and reasoning go well with one another. Reasoning differs fundamentally from revision, by drawing not on new information but on inferences from existing beliefs. For logicians, reasoning happens in language, and is a process of drawing conclusions from initially believed premises. Reasoning works much more naturally with categorical than with graded beliefs.

Rational choice theorists and logicians are both right in their own terms, since both models of belief fulfill the purpose set by the respective discipline. But can both kinds of belief coexist in the same agent? Such an agent would for instance simultaneously believe that it rains with subjective probability 2/3 and that it rains *simpliciter*. More generally, for any relevant proposition p, the

OPEN ACCESS

Edited by:
Rüdiger J. Seitz,
Heinrich Heine University of Düsseldorf, Germany

Reviewed by:
Hannes Leitgeb,
Ludwig Maximilian University of Munich, Germany
Lluis Oviedo,
Pontifical University Antonianum, Italy

***Correspondence:**
Franz Dietrich
fd@franzdietrich.net

Specialty section:
This article was submitted to Theoretical and Philosophical Psychology, a section of the journal Frontiers in Psychology

Received: *18 November 2021*
Accepted: *08 February 2022*
Published: *21 March 2022*

Citation:
Dietrich F (2022) Categorical Versus Graded Beliefs. Front. Psychol. 13:817940. doi: 10.3389/fpsyg.2022.817940

agent would hold some subjective probability of p and some yes/no belief about p. Depending on the context, the agent might draw either on their categorical beliefs or on their graded beliefs. In some contexts, the agent might reason logically with categorical beliefs, by drawing inferences from existing beliefs, thereby forming new beliefs. When learning information, the agent might on the one hand logically revise categorical beliefs, and on the other hand Bayes-revise graded beliefs. In decision-making contexts, the agent might either use a simple heuristic based on categorical beliefs and values, or use a more sophisticated decision rule (possibly the expected-utility rule) based on graded beliefs and values. In short, each type of belief would play a different functional role. Neither type would be redundant, since each type is tailored to its own role, and each type outperforms the other in its own area of application. Under this attractive division-of-labor picture, both belief types would be legitimate components of psychology.

But this ecumenical picture can only be maintained if the two belief types are mutually compatible in some sense, i.e., can coexist coherently. What exactly coherence amounts to is a question on its own. *Prima facie*, one would expect the agent to categorically believe propositions in which they have high degree of belief, and to categorically disbelieve propositions in which they have low degree of belief. This has come to be known as the "Lockean Thesis" (Foley, 2009). As we shall however see, this thesis leads straight into the "lottery paradox," from where an active literature unfolds about whether and how both belief types are co-tenable.

2. FROM THE LOTTERY PARADOX TO A GENERAL IMPOSSIBILITY THEOREM ABOUT COEXISTENCE OF BOTH BELIEF TYPES

Our notion of "can coexist" is normative, not positive. That is, we do not describe real agents, but ask whether an idealised agent—perhaps called a "rational" agent—can hold both belief types. The coexistence of both belief types is challenged by the well-known *lottery paradox* (Kyburg, 1961). This paradox starts from the Lockean Thesis—whereby one believes a proposition categorically if and only if one has a high enough degree of belief in it—and shows that this thesis generates a serious problem: even when graded beliefs are perfectly rational, i.e., obey probability theory, the corresponding set of categorical beliefs, formed *via* the Lockean Thesis, can be irrational, i.e., neither consistent nor deductively closed.

Why? In the lottery paradox, you are given a book of 100 pages. You know that exactly one page is black and all others are white. You have no idea about which page is black. So for each page you have a subjective probability of 99/100 that it is white. This subjective probability is high enough to make you (categorically) believe that the page is white. Meanwhile you have a subjective probability of 1 that not all pages are white. This maximal subjective probability is of course high enough to make you (categorically) believe that not all pages are white. Your categorical beliefs present two logical flaws. For one, you believe that the first page is white, that the second page is white, and so on, but you fail to believe an implication of these 100 beliefs, namely that all pages are white—a violation of deductive closure. Worse, you believe the opposite of this implication, namely that *not* all pages are white—a violation of logical consistency.

Though special in its setup, the lottery paradox highlights a deep and general problem. The literature has responded to it in different ways. One approach is "constructive" and consists in introducing, defending, or criticising concrete non-Lockean relations between both belief types that avoid the paradox. A number of potential relations are on the table; see for instance the "odds-threshold rule" in Lin and Kelly (2012a,b), the "stability theory" in Leitgeb (2014, 2017), and the "premise-based," "distance-based," "sequential," "relevance-based," and "holistic-threshold-based" relations in Dietrich and List (2018, 2021). Douven and Rott (2018) critically analyse the first two mentioned proposals. Earlier work about the lottery paradox includes (Hawthorne and Bovens, 1999; Douven and Williamson, 2006; Douven and Romeijn, 2007).

Taking an axiomatic rather than constructive approach, the lottery paradox was recently generalised into an impossibility theorem, proved in two versions by Dietrich and List (2018, 2021). Other impossibility theorems generalising the paradox were proved by Schurz (2019).

We here sketch Dietrich and List's theorem. It says: There is *no* form of coexistence of both belief types that respects certain initially plausible conditions. What are these conditions? There are six of them. The first three pertain each to one belief type only, and the next three pertain to the relationship between both belief types. Here are informal statements of the conditions:

1. The agent only ever holds categorical beliefs that are consistent and deductively closed.
2. The agent only ever holds degrees of belief that are probabilistically coherent (so that, for instance, the probability of "rain or snow" is the sum of the probabilities of "rain" and "snow.")
3. Any (probabilistically coherent) degrees of belief are allowed, i.e., can be held jointly with at least some categorical beliefs.
4. Whenever a proposition is believed with subjective probability 1, then it is believed categorically.
5. The two belief types impose at least some non-trivial constraints on one another, rather than being essentially independent of one another.
6. Any dependence between the two belief types is "local" ("proposition-wise") rather than 'global' ("holistic,") in a sense defined below. For instance, the Lockean Thesis postulates a purely local dependence, since the categorical belief in a proposition depends solely on the degree of belief in *this* proposition.

To state these conditions more precisely, let me sketch the formal setup. Consider a set X of propositions (or events) of interest; in the lottery paradox, X contains at least propositions about page colors. X could be very large, possibly containing *all* meaningful propositions, or very small, possibly containing only

propositions about a particular topic such as page colors, the Corona virus, or tomorrow's weather[1]. The agent's graded beliefs are represented by a *degree-of-belief function Pr* that assigns to each proposition $p \in X$ a subjective probability $Pr(p) \in [0, 1]$. The agent's categorical beliefs are represented by a *belief set* $B \subseteq X$, containing the (categorically) believed propositions. Certain combinations (Pr, B) of a degree-of-belief function and a belief set are "coherent" or "(rationally) co-tenable," the others are not. Formally, coherence or co-tenability defines a binary relation between degree-of-belief functions Pr and belief sets B—the relation of being mutually coherent or co-tenable.

The theorem assumes that this coherence relation satisfies six conditions. They were stated informally above. Here, are more formal re-statements:

1. Categorical beliefs are logically coherent: all permissible belief sets B are logically consistent and deductively closed. "Permissible" means that B is coherent with at least one degree-of-belief function Pr. Deductive closedness is defined relative to X: every proposition *from* X that B entails is contained in B.
2. Graded beliefs are probabilistically coherent: any permissible degree-of-belief function Pr obeys the laws of probability. "Permissible" means that Pr is coherent with at least one belief set B.
3. No coherent graded beliefs are ruled out: every probabilistically coherent degree-of-belief function Pr is permissible. "Permissible" was just defined.
4. Completely certain propositions are categorically believed: for any coherent combination (Pr, B) and any proposition $p \in X$, if $Pr(p) = 1$ then $p \in B$.
5. The two belief types are non-loosely related: at least one (permissible) degree-of-belief function Pr requires to believe some proposition $p \in X$ that is not completely certain, i.e., satisfies $Pr(p) \neq 1$. This rules out that all categorical beliefs are optional except under complete certainty. Technically, a degree-of-belief function Pr is said to "require" to believe a proposition p if p is contained in all belief sets coherent with Pr.
6. Any dependence between both belief types is "local" or "proposition-by-proposition:" whether the graded beliefs *require* to believe a given proposition only depends on the graded belief in *this* proposition (where "require to believe" was just defined). For instance, if the graded beliefs require to believe in rain, then changing the degree of belief in sunshine without changing the degree of belief in rain does not lift the requirement to believe in rain. The Lockean Thesis is an example of locality: here, believing a proposition is required if and only if the degree of belief in this proposition is high enough.

[1]Propositions could for instance be modeled as sets of possible worlds, i.e., subsets of some fixed underlying set Ω of possible worlds. This "semantic" or "set-theoretic" notion of proposition is common in rational-choice theory and probability theory, where propositions are usually called "events." Technically, X should be non-empty and closed under negation, i.e., a union of disjoint pairs of a proposition and its negation (The negation of a semantic proposition is of course its set-theoretic complement).

The impossibility theorem says: these six conditions are mutually incompatible[2].

A special kind of coherence relation deserves being mentioned: so-called *functional* or *deterministic* relations, in which the graded beliefs fully determine the categorical beliefs. Formally, functionally means that each permissible degree-of-belief function Pr is coherent with exactly one belief set B. Such a functional relation can be captured by a *binarization function* f which maps any (permissible) degree-of-belief function Pr to the corresponding belief set $B = f(Pr)$. The mentioned impossibility result was initially stated under the assumption of functionality, hence as a theorem about the inexistence of any binarization *function* satisfying certain conditions (Dietrich and List, 2018). To our later surprise, the impossibility extends to the much broader case without functionality assumption (Dietrich and List, 2021). The non-functional case allows one's categorical beliefs to be related much more loosely to one's graded beliefs: one's degrees of belief could impose almost no constraints on categorical beliefs, thereby leaving much freedom in what to believe categorically. Despite such freedom, it remains impossible to hold both belief types in accordance with the mentioned conditions.

3. WHAT TO MAKE OF THIS IMPOSSIBILITY?

Different reactions to the impossibility theorem are imaginable. Either one takes rational agents to have only graded beliefs, no categorical beliefs—against the logical paradigm. Or one takes rational agents to have only categorical beliefs, no graded beliefs—against the rational-choice-theoretic paradigm. Or one maintains coexistence, but gives up some of the conditions assumed in the incompatibility theorem. As a matter of fact, most conditions seem inescapable. But there are two important exceptions:

- One might give up the locality of the dependence between both belief types (condition 6). This would in particular give up the Lockean Thesis. Although locality is less demanding than the Lockean Thesis—it for instance does not imply functionality—locality is a strong constraint, normatively and mathematically, so that sacrificing it might be in order. Examples of non-local ("holistic") relations between both belief types are the mentioned relations in Lin and Kelly (2012a,b), Leitgeb (2014, 2017), or

[2]The theorem assumes that the set X of propositions under consideration contains enough interconnections. Unsurprisingly, there is no impossibility of holding both belief types relative to X if X contains only one (contingent) proposition-negation pair $p, \neg p$, or more generally if all such pairs in X are logically independent, because holding consistent and deductively closed belief sets is trivial for such X. However the impossibility result does for instance apply under the standard assumption that the set of propositions X forms a Boolean algebra that is not trivially small, i.e., contains more than one (contingent) proposition-negation pair. Boolean algebras are usually taken for granted. Interestingly, they are not essential for the theorem. For details, we refer to Dietrich and List (2021).

Dietrich and List (2018, 2021). Some of these relations are functional, others are non-functional.

- More radically, one could turn to a different theory of graded beliefs, by giving up probabilistic beliefs in favor of some other notion of graded belief. Multi-valued logic and ranking theory provide alternative kinds of graded belief. This intervention goes beyond giving up condition 2, since it alters the formal object of a degree of belief, and hence the range of degree-of-belief functions (initially, the set $[0, 1]$). Interestingly, ranking-theoretic beliefs (Spohn, 2012) would escape the impossibility and allow for a viable coexistence of graded and categorical beliefs—even a functional one. Needless to say, orthodox rational choice theory would be reluctant to replace "their" probabilistic paradigm by an altogether different, albeit graded, notion of belief.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

This work was supported by the French Research Agency through the Grants ANR-17-CE26-0003, ANR-16-FRAL-0010, and ANR-17-EURE-0001.

ACKNOWLEDGMENTS

The author thanks two referees and participants of the conference *Credition—An Interdisciplinary Challenge* (Hansson, 2021) for useful feedback.

REFERENCES

Dietrich, F., and List, C. (2018). From degrees of belief tobinary beliefs: lessons from judgment-aggregation theory. *J. Philos.* 115, 225–270. doi: 10.5840/jphil2018115516

Dietrich, F., and List, C. (2021). "The relation between degrees of belief and binary beliefs: a general impossibility theorem," in *Lotteries, Knowledge, and Rational Belief. Essays on the Lottery Paradox*, ed I. Douven (Cambridge: Cambridge University Press), 223–254.

Douven, I., and Romeijn, J.-W. (2007). The discursive dilemma as a lottery paradox. *Econ. Philos.* 23, 301–319. doi: 10.1017/S0266267107001502

Douven, I., and Rott, H. (2018). From probabilities to categorical beliefs: going beyond toy models. *J. Logic Comput.* 28, 1099–1124. doi: 10.1093/logcom/exy017

Douven, I., and Williamson, T. (2006). Generalizing the lottery paradox. *Brit. J. Philos. Sci.* 57, 755–779. doi: 10.1093/bjps/axl022

Foley, R. (2009). "Beliefs, degrees of belief, and the Lockean thesis," in *Degrees of Belief*, eds F. Huber, and C. Schmidt-Petri (Dordrecht: Springer), 37–47.

Hansson, S. O. (2021). A characterization of probability-based dichotomous belief revision. *Stud. Logica.* doi: 10.1007/s11225-021-09961-2

Hawthorne, J., and Bovens, L. (1999). The preface, the lottery, and the logic of belief. *Mind* 108, 241–264.

Kyburg, H. E., Jr. (1961). *Probability and the Logic of Rational Belief.* Middletown, CT: Wesleyan University Press.

Leitgeb, H. (2014). The stability theory of belief. *Philos. Rev.* 123, 131–171. doi: 10.1215/00318108-2400575

Leitgeb, H. (2017). *The Stability of Belief. How Rational Belief Coheres With Probability*. Oxford: Oxford University Press.

Lin, H., and Kelly, K. T. (2012a). Propositional reasoning that tracks probabilistic reasoning. *J. Philos. Logic* 41, 957–981. doi: 10.1007/s10992-012-9237-3

Lin, H., and Kelly, K. T. (2012b). A geo-logical solution to the lottery paradox. *Synthese* 186, 531–575. doi: 10.1007/s11229-011-9998-1

Schurz, G. (2019). Impossibility results for rational belief. *Noûs* 53, 134–159. doi: 10.1111/nous.12214

Spohn, Wolfgang (2012). *The Laws of Belief: Ranking Theory and Its Philosophical Applications*. Oxford: Oxford University Press.

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Dietrich. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Psychology

TYPE Opinion
PUBLISHED 11 August 2022
DOI 10.3389/fpsyg.2022.948330

Check for updates

OPEN ACCESS

EDITED BY
Rüdiger J. Seitz,
Heinrich Heine University of Düsseldorf, Germany

REVIEWED BY
David Kronemyer,
University of California, Los Angeles, United States

*CORRESPONDENCE
Andrea Vestrucci
avestrucci@sksm.edu

SPECIALTY SECTION
This article was submitted to Theoretical and Philosophical Psychology, a section of the journal Frontiers in Psychology

RECEIVED 19 May 2022
ACCEPTED 19 July 2022
PUBLISHED 11 August 2022

CITATION
Vestrucci A (2022) Deontic-doxastic belief revision and linear system model. *Front. Psychol.* 13:948330. doi: 10.3389/fpsyg.2022.948330

COPYRIGHT
© 2022 Vestrucci. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Deontic-doxastic belief revision and linear system model

Andrea Vestrucci[1,2]*

[1]Faculty of Information Systems and Applied Computer Sciences, University of Bamberg, Bavaria, Germany, [2]Starr King School, Oakland, CA, United States

KEYWORDS

belief revision, Dynamic Epistemic Logic, AGM theory, doxastic voluntarism, deontic logic, doxastic logic, belief models, logic of belief revision

Introduction

The article presents a doxastic-nested-deontic formalization of epistemic deontology (Feldman, 2000; Forrai, 2021) for static and dynamic belief revision, in AGM theory (and extensions) and Dynamic Epistemic Logic, respectively. The article also introduces a linear system model for beliefs[1].

Doxastic and deontic logics axiomatize propositions about beliefs ("it is believed that") and prescriptions ("it is obligatory that"), respectively. They belong to the family of modal logics.

Static and dynamic belief revisions follow from adding conflicting information to a belief database: in the static setting the doxastic value of the information is fixed (revision is non-iterated); in the dynamic setting information can be revised (revision can be iterated). In light of this, static belief revision might seem incompatible with belief update (Katsuno and Mendelzon, 1992) since update deals with information change (Seitz et al., 2018). This position has been variously challenged (Friedman and Halpern, 1994; Peppas and Williams, 1995; Gabbay, 1999; Aucher, 2004). Belief revision theories do relate to models for database update (Val and Shoham, 1994; Williams, 1997; Ditmarsch et al., 2008).

The article's outputs address: doxastic voluntarism; a paradox in strong epistemic deontology; the specificity of religious beliefs (Oviedo and Szocik, 2020).

1 The limits of this article do not allow consideration of other belief revision theories – e.g., ranking theory (Spohn, 1988, 2012; Huber, 2006, 2021) and Bayesian model (Brown et al., 2019) – nor discussion of AGM theory being an idealization of actual human doxastic agents in light of the logical, epistemological, and empirical simplifications involved in AGM (Wassermann, 1999; Berto, 2019). However, this idealization is useful to formalize belief revision (Hansson, 2022), thus paving the way to models more adherent to real doxastic situations, such as Dynamic Epistemic Logic (Section 3) and the linear system model (Section 4). Any adherence is nevertheless affected by the distinction between model and modeled object. AGM's clarity and logical and computational versatility (Delgrande et al., 2013; Spurkeland et al., 2013) make it a good candidate to introduce the doxastic-nested-deontic grammar in this article.

Static belief revision

Beliefs are elements of a set $\mathcal{B}$ (Alchourrón et al., 1985; Gärdenfors, 1988) over which three relations are defined: (1) logical consequence "$\vdash$" (Gärdenfors, 1984; Alchourrón et al., 1985); (2) epistemic entrenchment "$\preccurlyeq$" (Gärdenfors and Makinson, 1988), and (3) spheres inclusion "$\geq$" (Grove, 1988).

Concerning 1, elements of $\mathcal{B}$ are logical consequences of other elements (e.g., believing that tomorrow will rain follows from believing in the reliability of weather forecasts).

$$\mathcal{B} = Cn(\mathcal{B}) = \{\alpha : \mathcal{B} \vdash \alpha\}$$ (Huber, 2013; Hansson, 2022).

In case of a new information ϑ contradicting some elements of $\mathcal{B}$, $\mathcal{B}$ is revised ("$*$") by clearing $\mathcal{B}$ from all elements contradicted by ϑ, and adding ϑ:

$$\mathcal{B} * \vartheta = Cn(\mathcal{B} - \neg\vartheta \cup \{\vartheta\})$$ (Levi, 1977).

Concerning 2, entrenchment is a preorder on $\mathcal{B}$ (Peppas, 2008) based on belief firmness: the more a belief is entrenched, the more it costs to give it up. This applies also to logical consequences; thus 1 and 2 are related: $\alpha \vdash \beta \rightarrow \alpha \preccurlyeq \beta$ (Dominance postulate). Belief revision deals with clearing $\mathcal{B}$ from anything that is less or equally entrenched than all elements contradicted by ϑ, and adding ϑ.

$$\mathcal{B} * \vartheta = Cn(\{\psi \in \mathcal{B} : \neg\vartheta \prec \psi\} \cup \{\vartheta\}).$$

Revision consists in the "minimal mutilation" (Rott, 2000; Leitgeb, 2010) of $\mathcal{B}$ [keeping as much old beliefs as possible (Ditmarsch et al., 2008)], and the addition of ϑ.

Concerning 3, worlds w in which elements from $\mathcal{B}$ are true are placed on spheres ordered per inclusion. Given a Kripke model M, $[\mathcal{B}]_M = \{w \in W^M : M, w \vDash \varphi \; \forall\varphi \in \mathcal{B}\}$. Inclusion can be grasped as *plausibility* order (Peppas, 2008): the most plausible possible worlds are located on spheres with the least radius. Thus, 3 is related to 2: $\varphi \preccurlyeq \psi \cong [\varphi]_M \geq [\psi]_M$. Considering $[\vartheta]_M = \{w \in W^M : M, w \vDash \vartheta\}$, agent a's belief in φ conditioned on ϑ ("$B_a^\vartheta\varphi$") is true in the minimal-radius spheres (i.e., most plausible worlds) in which ϑ is also true. By simplifying Baltag and Renne, 2016:

$$M, w \vDash B_a^\vartheta\varphi \in \mathcal{B} * \vartheta \equiv min_a([\vartheta]_M) \subseteq [\varphi]_M.$$

This corresponds to making ϕ a *safe* belief (Baltag et al., 2008):

$$M, w \vDash B_a^\vartheta\varphi \in \mathcal{B}*\vartheta \equiv M, w \vDash \Box_a\varphi.$$

Thus, the formula for epistemic deontology of belief revision in a static setting corresponds to:

$$((\Box_a\vartheta \wedge \exists x \in \mathcal{B}_a : \vartheta \nvdash x) \rightarrow) \; \Diamond\mathbf{O}(\forall\phi \in \mathcal{B}_a, \; \phi \rightarrow \Box_a\varphi).$$

Since $B_a^\vartheta\varphi$ is a doxastic conditional, $\Diamond\mathbf{O}$ is a conditioned obligation presupposing that agent a has at least a safe belief on ϑ, and that ϑ is in a negative relation with at least one element of $\mathcal{B}$. The formula represents the duty to increase the epistemic degree of set $\mathcal{B}$: it formalizes Kant's "*Sapere aude!*" (Kant, 2013).

The nested formula applies to negative doxastic voluntarism (NDV), the idea that we have control not over belief formation but over belief withdrawal (Rott, 2017). The formula translates "belief withdrawal" into "epistemic-degree-increase duty," and it associates the notion of "negative control" to the whole spectrum of duty realizations, including duty *non*-realization; thus, voluntarism pertains also to the refusal of epistemic degree increase. Moreover, since the duty is conditional, the formula expands NDV to include (or even presuppose) belief expansion ($\Box_a\vartheta$).

Dynamic belief revision

In static belief revision, information ϑ is included in the revised set (Success postulate: $\vartheta \in \mathcal{B} * \vartheta$). Thus, static revision assumes the epistemic value of ϑ to be unchangeable. This is problematic, e.g., in the case of Moore sentences involving higher-order beliefs (Baltag et al., 2008). To amend this, ϑ shall be considered susceptible of revision too. Research in dynamic belief revision distinguishes at least three epistemic degrees of ϑ (van Benthem, 2007; Baltag and Smets, 2009; Baltag et al., 2014): (1) ϑ is "hard information" issued from an infallible source: it is neither revisable nor revocable; (2) ϑ is "soft information" from a fallible, yet highly reliable source; (3) ϑ is "soft information" from a barely trusted source (truthfulness can be easily given up).

To these three doxastic degrees correspond three types of dynamic belief revision:

1. Radical revision $[!\vartheta]$: it eliminates all $\neg\vartheta$-worlds and the previous plausibility order is preserved between the remaining worlds.
2. Lexicographic (radical) revision $[\Uparrow\vartheta]$: all ϑ-worlds are made more plausible than $\neg\vartheta$-words, and the rest of the order is unchanged.
3. Conservative (neutral) revision $[\uparrow\vartheta]$: the most plausible ϑ-words become the most plausible worlds overall, and all rest is unchanged.

Thus, the formula for the epistemic deontology of belief revision in a dynamic setting corresponds to (the lexicographic formula; van Benthem, 2011 is a generalization of the conservative one):

$$\Box/\Diamond\mathbf{O}\begin{pmatrix}[!\vartheta]B_a\varphi \equiv (\vartheta \rightarrow B_a^\vartheta([!\vartheta]\varphi)) \\ M, w \vDash [\Uparrow\vartheta]B_a\varphi \equiv M \Uparrow \vartheta, w \vDash B_a\varphi \\ [\uparrow\vartheta]B_a\varphi\end{pmatrix}$$

The deontic operator might be not conditioned since the doxastic conditions for revision are *within* the obligation. This would introduce to a strong epistemic deontology: under no

condition a belief is allowed to be held if no sufficient evidence supports it.

This leads to a paradox in strong epistemic deontology. Let's assume two scenarios: 1. the revision process halts, 2. it does not halt. In 1, the revision halts because a belief has received sufficient evidence to be no longer revisable. Thus, it is not even a (safe) belief: it is infallible and indefeasible *knowledge* resisting any information (true or false) (Baltag and Smets, 2008). In 2, the reiterated halt delay means that the collection of evidence never ends: the belief is never allowed to be held. Thus, from a strong epistemic deontology, no belief is ever legitimate, regardless of the doxastic degree of it: either a belief is transformed into knowledge, or it is never sufficiently justified. Hence, the paradox: believing is always wrong *for the fact itself of believing*.

The deontic encapsulation of dynamic belief revision might address this paradox by including not only belief revision, but also information ϑ in the *deontic* environment. The duty of belief revision is not unconditioned, but conditioned by the duty of evaluating the object itself of duty (collecting ϑ) either positively or negatively. This might include the rejection or neglection of information ϑ as forms of epistemic deontology satisfaction.

Linear system model

The aforementioned theories conceive beliefs as elements of a set. This set model imposes at least three requirements: (1) The elements of a belief set must be orderable according to some (pre)orders; (2) The belief set must be somehow coherent, and belief revision corresponds to the maximal preservation of this coherency; (3) A new information is needed, which contradicts at least one belief. The weight of the revision work is proportional to the number of beliefs connected to the new information, and to the negativity of this connection.

Do these three requirements apply to all belief set revisions? If we take the case of religious beliefs, then: (1) An ordering relation implies a comparability between the set elements which at its turn implies a homogeneity of the elements' epistemic bases. However, religious beliefs cover different epistemic spheres: metaphysical, moral, aesthetic, pragmatic, etc. It's not clear how beliefs referring to such different epistemic spheres can be fully comparable. (2) The issue of theodicy is evidence against the (at least *prima facie*) coherency of religious beliefs since theodicy tries to address the incompatibility between the belief in divine omnipotence and the belief in divine justice. (3) The revision of a religious set may start not only from external information, but also from introspection, i.e., the internal evaluation and investigation of one's faith.

Thus, I propose an alternative model, in which beliefs are elements of a system of linear equations. This linear system model has at least two advantages compared to the set model:

1. Bottom-up organization. In the set model, a belief's relevance depends on its being an element of a set, i.e., the belief characteristics are deduced from the set definition. This is why in the set model beliefs constitute a coherent unity and are ordainable: conditions 1 and 2 follow from the application of the set model to beliefs (and their revision). In the linear system model, the solution of the system is given by the linear equations (the elements) constituting the system. Thus, the belief's characteristics precede (and not follow from) the system including them: rather than selecting beliefs in light of a certain model (a certain definition of belief set), the model is constructed and constantly readjusted in light of the elements we aim to investigate. This bottom-up organization respects the epistemic "matter" by building the model *upon* this matter.
2. Representation of belief *stratification*. Beliefs are stratified vertically and horizontally. The vertical stratification is the succession of beliefs, represented by the order of the equations in the linear system; this succession is not necessarily a preorder since the equations' order does not change the system's solution. However, the vertical stratification has a *procedural* function: it eases the substitution of the variables that are gradually known. Moreover, the system might allocate different epistemic spheres in different vertical strata, thus not overlapping epistemically distinct beliefs. The horizontal stratification is the composition of a belief as a sum of *sub-beliefs*: each sub-belief is a part of the greater belief, and their order in the summation corresponds to their relevance within the whole belief. For example, the belief in the 10 commandments is composed by the sub-beliefs in all single commandments, each sub-belief doxastically introducing the successive.

This linear system model is:

$$R = \begin{cases} a_{1,1}\Box_1 + a_{1,2}\Box_2 + a_{1,3}\Box_3 + \ldots + a_{1,n}\Box_n = w_1 \\ \ldots \\ a_{m,1}\Box_1 + a_{m,2}\Box_2 + a_{m,3}\Box_3 + \ldots + a_{m,n}\Box_n = w_m \end{cases}$$

$$R = \bigcap_{i=1}^{m} \sum_{j=1}^{n} a_{i,j}\Box_j = \bigcap_{i=1}^{m} w_i$$

A system R represents the vertical stratification of m beliefs: it is the intersection of m polynomial equations with n variables. In each equation, the coefficient $a_{i,j}$ is the content of a belief or sub-belief, e.g., an equation with ten coefficients might represent the belief in the ten commandments. The variable $\Box_j$ is the doxastic value associated to the belief content in position j. The doxastic value is the same for all coefficients in position j since it follows the horizontal stratification: a sub-belief in j is the doxastic "step" to reach the sub-beliefs in positions $k > j$. The constant term w_i expresses the possible world plausibility of the entire polynomial.

This model permits a more "economic" belief revision. In the set model (for both static and dynamic scenarios), revision consists in a modification of the set structure: a subset is eliminated or displaced in light of new information. In the linear system model, the elimination of a belief (equation) does not necessarily affect the system: the condition is for the number of equations to be at least equal to the number of doxastic values; one can obtain this by readjusting some coefficients (sub-beliefs), e.g., expunging the *filioque* belief without touching other religious-metaphysical beliefs, but maybe modifying some religious-aesthetic beliefs.

This model also permits a simpler procedure to compare different belief systems. For example, an orthodox and a non-orthodox Christian might have belief systems (resp. R_1 and R_2) which differ for the third equation (non-*filioque* in R_1, in bold), but are identical for the rest.

$$R_1 = \begin{cases} a_{1,1}\square_1 + a_{1,2}\square_2 + a_{1,3}\square_3 + \ldots + a_{1,10}\square_{10} = w_1 \\ b_{2,1}\square_1 = w_2 \\ \mathbf{c_{3,1}\square_1 + 0\square_2 + c_{3,3}\square_3 = w_3} \\ \ldots \end{cases}$$

$$R_2 = \begin{cases} a_{1,1}\square_1 + a_{1,2}\square_2 + a_{1,3}\square_3 + \ldots + a_{1,10}\square_{10} = w_1 \\ b_{2,1}\square_1 = w_2 \\ \mathbf{c_{3,1}\square_1 + c_{3,2}\square_2 + c_{3,3}\square_3 = w_3} \\ \ldots \end{cases}$$

Matrix form is even clearer:

$$R_1 = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & a_{1,10} \\ b_{2,1} & 0 & 0 & \cdots & 0 \\ \mathbf{c_{3,1}} & \mathbf{0} & \mathbf{c_{3,3}} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

$$R_2 = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & a_{1,10} \\ b_{2,1} & 0 & 0 & \cdots & 0 \\ \mathbf{c_{3,1}} & \mathbf{c_{3,2}} & \mathbf{c_{3,3}} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

The linear system model presents an intuitive approach to synthetically grasp a relationship between belief systems. Thus, the model might better capture the limits and extent of ecumenical and interreligious dialogues.

Discussion

Aspects of future investigation include: establishing the doxastic-nested-deontic grammar; assessing the approach it provides to doxastic voluntarism; presenting a deontic investigation of the epistemic deontology paradox; deepening the potentialities and weaknesses of the linear system model for beliefs; exploring belief translatability from the linear model to the set model and vice-versa.

Author contributions

The author confirms being the sole contributor of this work and has approved it for publication.

Funding

This paper was funded by Dr. Rüdiger Seitz, via the Volkswagen Foundation, Siemens Healthineers, and the Betz Foundation.

Acknowledgments

I thank Martin Aleksandrov, Christoph Benzmüller, Alessio Pirastu, and the reviewer for valuable comments on a previous version of this article.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Alchourrón, C., Gärdenfors, P., and Makinson, D. (1985). On the logic of theory change: partial meet contraction and revision functions. *J. Symbolic Logic* 50, 510–530. doi: 10.2307/2274239

Aucher, G. (2004). "A combined system for update logic and belief revision," in *Pacific Rim International Workshop on Multi-Agents* (Berlin: Springer), 1–17.

Baltag, A., Fiutek, V., and Smets, S. (2014). "DDL as an "Internalization" of dynamic belief revision," in *Krister Segerberg on Logic of Action*, ed R. Trypuz (Berlin: Springer), 253–280.

Baltag, A., and Renne, B. (2016). "Dynamic epistemic logic," in *The Stanford Encyclopedia of Philosophy*, Eds E. N. Zalta. Available online at: https://plato.stanford.edu/archives/win2016/entries/dynamic-epistemic/ (accessed July 26, 2022).

Baltag, A., and Smets, S. (2008). "A qualitative theory of dynamic interactive belief revision," in *TLG 3: Logic and the Foundations of Game and Decision Theory*, Eds G. Bonanno, W. van der Hoek, and M. Wooldridge (Amsterdam: Amsterdam University Press), 11–58.

Baltag, A., and Smets, S. (2009). "Group belief dynamics under iterated revision: fixed points and cycles of joint upgrades," in *Proceedings of Theoretical Aspects of Rationality and Knowledge TARK* (New York: Association for Computing Machinery), 41–50.

Baltag, A., van Ditmarsch, H. P., and Moss, L. S. (2008). "Epistemic logic and information update," in *Philosophy of Information*, Eds P. Adriaans and J. van Benthem (Cambridge, MA: MIT Press), 361–456.

Berto, F. (2019). Simple hyperintensional belief revision. *Erkenntnis* 84, 559–575. doi: 10.1007/s10670-018-9971-1

Brown, W., Gyenis, Z., and Rédei, M. (2019). The modal logic of bayesian belief revision. *J. Philos. Logic* 48, 809–824. doi: 10.1007/s10992-018-9495-9

Delgrande, J., Peppas, P., and Woltran, S. (2013). "AGM-style belief revision of logic programs under answer set semantics," in *Logic Programming and Nonmonotonic Reasoning. LPNMR 2013*, Eds P. Cabalar and T. C. Son (Berlin: Springer), 264–276.

Ditmarsch, H., Hoek, W., and Kooi, B. (2008). *Dynamic Epistemic Logic*. Berlin: Springer.

Feldman, R. (2000). The ethics of belief. *Philos. Phenomenol. Res.* 60, 667–695. doi: 10.2307/2653823

Forrai, G. (2021). Doxastic deontology and cognitive competence. *Erkenntnis* 86, 687–714. doi: 10.1007/s10670-019-00126-1

Friedman, N., and Halpern, J. Y. (1994). "A knowledge-based framework for belief change, part II: revision and update," in *Proceedings of the Fourth International Conference on Principles of Knowledge Representation and Reasoning (KR'94)* (San Francisco, CA, USA: Morgan Kaufmann Publishers), 190–201.

Gabbay, D. M. (1999). "Compromise update and revision: a position paper," in *Dynamic Worlds*, eds R. Pareschi and B. Fronhöfer (Dordrecht: Springer), 111–148.

Gärdenfors, P. (1984). Epistemic importance and minimal changes of belief. *Australas. J. Philos.* 62, 136–157. doi: 10.1080/00048408412341331

Gärdenfors, P. (1988). *Knowledge in Flux: Modeling the Dynamics of Epistemic States*. Cambridge, MA: The MIT press.

Gärdenfors, P., and Makinson, D. (1988). "Revisions of knowledge systems using epistemic entrenchment," in *Proceedings of Theoretical Aspects of Reasoning About Knowledge* (San Francisco (CA): Morgan Kaufmann Publishers), 83–95.

Grove, A. (1988). Two modellings for theory change. *J. Philos. Logic* 17,157–170. doi: 10.1007/BF00247909

Hansson, S. O. (2022). "Logic of belief revision," in *The Stanford Encyclopedia of Philosophy*, Ed E. N. Zalta. Available online at: https://plato.stanford.edu/archives/spr2022/entries/logic-belief-revision/ (accessed July 26, 2022).

Huber, F. (2006). Ranking functions and rankings on languages. *Artif. Intell.* 170, 462–471. doi: 10.1016/j.artint.2005.10.016

Huber, F. (2013). Belief revision I: the AGM theory. *Philos. Compass* 8, 604–612. doi: 10.1111/phc3.12048

Huber, F. (2021). *Belief and Counterfactuals: A Study in Means-End Philosophy*. Oxford: Oxford University Press.

Kant, I. (2013). *An Answer to the Question: "What is Enlightenment?"*. London: Penguin Books.

Katsuno, H., and Mendelzon, A. (1992). "On the difference between updating a knowledge base and revising it," in *Belief Revision*, Ed P. Gärdenfors (Cambridge: Cambridge University Press), 183–203.

Leitgeb, H. (2010). On the Ramsey Test without triviality. *Notre Dame J. Formal Logic* 51, 21–54. doi: 10.1215/00294527-2010-003

Levi, I. (1977). Subjunctives, dispositions and chances. *Synthese* 34, 423–455. doi: 10.1007/BF00485649

Oviedo, L., and Szocik, K. (2020). Religious – And Other Beliefs: How Much Specificity? *Sage Open*. [Epub ahead of print].

Peppas, P. (2008). "Belief revision," in *Handbook of Knowledge Representation,* Eds F. van Harmelen, V. Lifschitz, and B. Porter (Amsterdam: Elsevier), 317–359.

Peppas, P., and Williams, M. A. (1995). Constructive modelings for theory change. *Notre Dame J. Formal Logic* 36, 120–133. doi: 10.1305/ndjfl/1040308831

Rott, H. (2000). Two dogmas of belief revision. *J. Philos.* 97, 503–522. doi: 10.2307/2678489

Rott, H. (2017). Negative Doxastic Voluntarism and the concept of belief. *Synthese* 194, 2695–2720. doi: 10.1007/s11229-016-1032-1

Seitz, R. J., Paloutzian, R. F., and Angel, H. F. (2018). From believing to belief: a general theoretical model. *J. Cogn. Neurosci.* 30, 1254–1264. doi: 10.1162/jocn_a_01292

Spohn, W. (1988). "Ordinal conditional functions: a dynamic theory of epistemic states," in *Causation in Decision, Belief Change, and Statistics II*, Eds W. L. Harper and B. Skyrms (Dordrecht: Kluwer), 105–134.

Spohn, W. (2012). *The Laws of Belief: Ranking Theory and Its Philosophical Applications*. Oxford: Oxford University Press.

Spurkeland, J. S., Jensen, A. S., and Villadsen, J. (2013). Belief revision in the GOAL agent programming language. *Int. Scholarly Res. Notices* 2013:632319. doi: 10.1155/2013/632319

Val, A. D., and Shoham, Y. (1994). A unified view of belief revision and update. *J. Logic Comput.* 4, 797–810. doi: 10.1093/logcom/4.5.797

van Benthem, J. (2007). Dynamic logic for belief revision. *J. Appl. Non Classical Logics* 17, 129–155. doi: 10.3166/jancl.17.129-155

van Benthem, J. (2011). *Logical Dynamics of Information and Interaction*. Cambridge: Cambridge University Press.

Wassermann, R. (1999). Resource bounded belief revision. *Erkenntnis* 50, 429–446. doi: 10.1023/A:1005565603303

Williams, M.-A. (1997). "Belief revision as database update," in *Proceedings Intelligent Information Systems*, 410–414.

frontiers | Frontiers in Psychology

OPINION
published: 26 July 2022
doi: 10.3389/fpsyg.2022.947664

The Nature of Belief From a Philosophical Perspective, With Theoretical and Methodological Implications for Psychology and Cognitive Science

*Eric Schwitzgebel**

Department of Philosophy, University of California, Riverside, Riverside, CA, United States

Keywords: belief, representation, questionnaires, implicit bias, developmental psychology, philosophy, dispositions, behaviorism

INTRODUCTION

In recent academic philosophy, *representationalism* is probably the dominant model of belief. I favor a competing model, *dispositionalism*. I will briefly describe these views and their contrasting implications, including some theoretical and methodological implications relevant to research psychologists and cognitive scientists.

REPRESENTATIONALISM VS. DISPOSITIONALISM, DEFINITIONS

According to representationalism, to believe some proposition P (for example, that there's beer in the fridge or that men and women are intellectually equal) is to have a representation with the content P stored in your mind, available to be deployed in relevant reasoning. It's somewhat unclear how literally the "storage" idea is to be taken, but leading representationalists, such as Fodor and Mandelbaum (Fodor, 1987; Mandelbaum, 2014; Quilty-Dunn and Mandelbaum, 2018; Bendaña and Mandelbaum, 2021), appear to take the storage idea rather literally. One might compare to the concept of the "long-term memory store" in theories of memory. The stored representation counts as *available to be deployed in relevant reasoning* if it can be accessed when relevant. If asked whether men and women differ in intelligence, you'll retrieve the representation that men and women are intellectually equal, engage in some simple theoretical reasoning, and answer "no" (if you want to be honest, etc.). If you feel like drinking a cold beer, you'll retrieve the representation that beer is in the fridge, engage in some simple practical reasoning, and walk toward the kitchen to get the beer.

According to dispositionalism, to believe that P is to be disposed to act and react in ways that are characteristic of believers-that-P. Maybe there's a representation really stored in there; maybe not. If you are disposed to go to the fridge when you want a beer, if you are disposed to say "yes" when asked whether there's beer in the fridge, if you display surprise upon opening the fridge and finding no beer, etc., then you count as believing that there's beer in the fridge, regardless what underlying cognitive architecture enables this. Dispositionalism has its roots in philosophical behaviorism and Ryle (1949). However, I and other recent dispositionalists eschew behaviorism, allowing that some of the relevant dispositions can be "phenomenal" (i.e., pertaining to conscious experience), such as the disposition to *feel* (and not just exhibit) surprise upon opening the fridge and seeing no beer, and other dispositions can be cognitive (i.e., pertaining to inference or other cognitive transitions), such as the disposition to *draw the conclusion* that there is beer in the house (Schwitzgebel, 2002, 2021).

OPEN ACCESS

Edited by:
Rüdiger J. Seitz, Heinrich Heine University of Düsseldorf, Germany

Reviewed by:
Dimitrios Stamovlasis, Aristotle University of Thessaloniki, Greece

***Correspondence:**
Eric Schwitzgebel
eschwitz@ucr.edu

Specialty section:
This article was submitted to Theoretical and Philosophical Psychology, a section of the journal Frontiers in Psychology

Received: *19 May 2022*
Accepted: *22 June 2022*
Published: *26 July 2022*

Citation:
Schwitzgebel E (2022) The Nature of Belief From a Philosophical Perspective, With Theoretical and Methodological Implications for Psychology and Cognitive Science. Front. Psychol. 13:947664. doi: 10.3389/fpsyg.2022.947664

Representationalism commits to a particular type of cognitive architecture—the storage of representational contents matching the contents of the believed propositions—and it is to a substantial extent neutral about the extent to which the stored contents are behavior-guiding. Dispositionalism commits to belief as behavior-guiding, while remaining neutral on the underlying architecture. The difference matters to psychological theory and method as I will now explain.

IN-BETWEEN BELIEVING

On representationalism, it's natural to think of belief as a yes/no matter. P is either stored or it's not. You either believe it or you don't. Representations can't normally be "half-stored." What would that even mean? If the representation isn't retrieved when relevant, it's a "performance" failure; the underlying "competence" is still there, as long as it could in principle be retrieved in some circumstances. This leads some representationalists, especially Mandelbaum, to unintuitive views about what we believe. For example, if someone tells you "dogs are made of paper," Mandelbaum holds that you will believe that proposition—even after you reject it as obviously false—because the representation gets stored and starts influencing your cognition. Of course you also simultaneously believe that dogs are not made of paper.

On dispositionalism, believing is more like having a personality trait: You match the dispositional profile to some degree, just like you might match the dispositional profile characteristic of extraversion to some degree. Sometimes, the match might be nearly perfect. I might have all the dispositions characteristic of the belief that there's beer in my fridge. Other times, the match might be far from perfect. Cases of highly imperfect match can be described as *in-between* cases of belief.

Consider the belief that men and women are intellectually equal. Someone—call him the "implicit sexist"—might be disposed to act and react in some ways that are characteristic of that belief. He might say "men and women are intellectually equal" with a feeling of confidence and sincerity, ready to defend that view passionately in a debate. Other dispositions might tilt the other way. He might feel surprised if a woman makes an intelligent comment at a meeting, and it might take more evidence to convince him that a women is smart than that a man is smart.

Or consider gradual forgetting. In college, I knew the last name of my roommate's best friend. I could easily recall it. Over time, as memory faded, I would have been able to recognize it, picking it out from nearby alternatives, but recall would have been weaker. As memory continued to fade, I would have recognized it less and less reliably until eventually it was utterly forgotten. During the intermediate phase, I would in some respects act and react like someone would believed his name was (let's say) Guericke, in other respects not. There was no precise moment at which the belief dropped from my mind, instead a long period of gradual, fading in-betweenness.

Dispositionalist views naturally invite us see belief as permitting in-between cases, as personality traits do. Representationalist views have more difficulty accommodating this idea.

CONTRADICTORY BELIEF

Conversely, representationalist views naturally allow for contradictory belief, as discussed in the "dogs are made of paper" example, while dispositionalist views appear to disallow the possibility of having contradictory beliefs. There seems to be no problem in principle in storing both the representation "P" and the representation "not-P." But one cannot simultaneously have the dispositional structure characteristic of believing that men and women are intellectually equal and the dispositional structure characteristic of believing that women are intellectually inferior. That would be like having the dispositional structure of an extravert and simultaneously the dispositional structure of an introvert—structurally impossible.

Given an implicit sexism case, then, representationalism tends to favor the idea that the sexist believes both that women and men are intellectually equal and that women are intellectually inferior. The two contradictory beliefs are both stored and accessible (perhaps in different cognitive subsystems, retrieved under different conditions). Dispositionalism tends to favor treating such cases as in-between cases of belief. Similarly for other inconsistent or conflicting attitudes: the Sunday theist/weekday atheist; the self-deceived husband who sincerely denies that his wife is cheating on him but sometimes acts as if he knows; the person who would say the road runs north-south if queried in one way but who would say it runs east-west if queried in another way.

Let me briefly defend the dispositionalist stance on this issue. We have no need for contradictory belief. It helps none to say of the implicit sexist that he believes both "men and women are intellectually equal" and "women are intellectually inferior." To make such a claim comprehensible, we need to present the details: In these respects he acts and reacts like an egalitarian, in these other respects he acts and reacts like a sexist. But now we've just given the dispositional characterization. If necessary—if there are good enough architectural grounds for it—we *might* still say that he has contradictory representations. But representation is not belief.

EXPLANATORY DEPTH VS. EXPLANATORY SUPERFICIALITY

Quilty-Dunn and Mandelbaum (2018) argue that representationalism has an explanatory depth that coheres well with the aims of cognitive science. If the belief that P is a relation to a stored representational content "P," we can explain how beliefs cause behavior (retrieving the stored representation does the causal work), we can explain why there's usually such a nice parallel between what we can say and what we can believe (speech and belief involve accessing the same pool of representations), and so forth. The dispositionalist approach, in contrast, is superficial: It points to the dispositional patterns but

it does not attempt to explain the causal mechanisms beneath those patterns.

While explanatory depth is a virtue *when available,* it is not a virtue in this particular case. To think that belief that P always, or typically, involves having an internal representational content "P" is a best empirically unsupported. (Contrast with the empirically well supported claim that the visual system represents motion in regions of the visual field.) At worst, it is a simplistic cartoon sketch of the mind. It's as if someone insisted that having the personality trait of extraversion required having an internal switch flipped to "E," because otherwise we'd be stuck without an internal causal explanation of extraverted patterns of behavior. Of course there are internal structures that help explain people's extraverted behavior, and of course there are internal structures that help explain people's implicitly sexist behavior and their beer-fetching behavior. But we need not define belief in terms of a simplistic representationalist understanding of those internal structures.

Still, a partial compromise is possible. It *might* be the case that internal representations of P are present whenever one believes that P. The dispositionalist need not deny this—any more than a personality theorist need not deny that extraversion might involve an heretofore-undiscovered E switch. The dispositionalist just doesn't define belief in terms of such structures, permitting a skeptical neutrality about them.

INTELLECTUALISM VS. PRAGMATISM

I will now introduce a second philosophical distinction. According to *intellectualism* about belief, sincere assent or assertion is sufficient or nearly sufficient for belief. According to *pragmatism* about belief, to really, fully believe you need not just to be ready to say P; you need also to act accordingly.

The intellectualism/pragmatism distinction cross-cuts the representationalism/dispositionalism distinction. However, I submit that the most attractive form of dispositionalism is also pragmatist. To really, fully believe that women are intellectually equal requires more than simply readiness to say they are. It requires not being surprised when a women makes an intelligent remark. It requires treating the women you encounter as if they are just as smart as men in the same circumstances. Alternatively, to really believe that your children's happiness is more important than their academic success it's insufficient to be disposed to *say* that is the case; you must also to live that way.

THE PROBLEM WITH QUESTIONNAIRES

I conclude with two methodological implications.

First, if pragmatist dispositionalism is correct, then you might not know what you believe. Do you really believe that men and women are intellectually equal? Do you really believe that your children's happiness is more important than their academic success? You'll *say* yes and yes. But how do you really live your life? You might be more in-betweenish than you think.

When psychologists want to explore broad, life involving beliefs and values, they often employ questionnaires. Questionnaires are easy! But if pragmatist dispositionalism is correct, questionnaires risk being misleading when asking about beliefs or other attitudes with an important lived component that can diverge from verbal endorsement. Questionnaires get at what you say, not at how you generally act.

A brief example: The Short Schwartz's Values Survey (Lindeman and Verkasalo, 2005) asks participants how important it is to them to achieve "power (social power, authority, wealth)" and various other goods. If intellectualism is the right way to think about values, this is an excellent methodology. However, if pragmatism is better, it's reasonable to doubt how well people know this about themselves.

DEVELOPING BELIEFS

Developmental psychologists often debate the age children reach various cognitive milestones, such as knowing that objects continue to exist even when they aren't being perceived and knowing that people can have false beliefs. If representationalism is correct, then it's natural to suppose that there is in fact some particular age at which each individual child finally comes to store the relevant representational content. However, if dispositionalism is correct, gradualism is probably more attractive: Such broad beliefs are slowly constructed, involving many relevant dispositions, which might accrete unevenly and unstably over months or years.

In my experience, developmental psychologists often endorse gradualism when explicitly asked. Yet their critiques of each other seem sometimes implicitly to assume the contrary. "Boosters" (who claim that knowledge in some domain tends to come early) reject as too demanding methodologies that appear to reveal later knowledge. "Scoffers" (who claim that knowledge in some domain tends to come late) reject as too easy methodologies that appear to reveal earlier knowledge. Each trusts only the methods that reveal knowledge at the "right" age. But while of course some methodologies might be flawed, the gradualist dispositionalist ought to positively expect that *across a variety of equally good methods* for discovering whether the child knows P, some should reveal much earlier knowledge than others, though none are flawed—because knowing that P is not a yes-or-no, not an on-or-off thing. There need be no one right age or set of methods. (For more on this issue, see Schwitzgebel, 1999.)

AUTHOR CONTRIBUTIONS

ES is solely responsible for all aspects of the manuscript.

FUNDING

This article is funded by Rüdiger Seitz, *via* the Volkswagen Foundation, Siemens Healthineers, and the Betz Foundation. The funders were not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

REFERENCES

Bendaña, J., and Mandelbaum, E. (2021). "The fragmentation of belief," in *The Fragmented Mind*, eds C. Borgoni, D. Kindermann, and A. Onofri (Oxford: Oxford University Press). doi: 10.1093/oso/9780198850670.003.0004

Fodor, J. A. (1987). *Psychosemantics*. Cambridge, MA: MIT Press. doi: 10.7551/mitpress/5684.001.0001

Lindeman, M., and Verkasalo, M. (2005). Measuring values with the Short Schwartz's Value Survey. *J. Person. Assess.* 85, 170–178. doi: 10.1207/s15327752jpa8502_09

Mandelbaum, E. (2014). Thinking is believing. *Inquiry* 57, 55–96. doi: 10.1080/0020174X.2014.858417

Quilty-Dunn, J., and Mandelbaum, E. (2018). Against dispositionalism: belief in cognitive science. *Philos. Stud.* 175, 2353–2372. doi: 10.1007/s11098-017-0962-x

Ryle, G. (1949). *The Concept of Mind*. New York, NY: Barnes and Noble.

Schwitzgebel, E. (1999). Gradual belief change in children. *Hum. Dev.* 42, 283–296. doi: 10.1159/000022637

Schwitzgebel, E. (2002). A phenomenal, dispositional account of belief. *Nous* 36, 249–275. doi: 10.1111/1468-0068.00370

Schwitzgebel, E. (2021). "The pragmatic metaphysics of belief," in *The Fragmented Mind*, eds C. Borgoni, D. Kindermann, and A. Onofri (Oxford: Oxford University Press). doi: 10.1093/oso/9780198850670.003.0015

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Schwitzgebel. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Psychology

OPINION
published: 29 June 2022
doi: 10.3389/fpsyg.2022.939949

Two Concepts of Belief Strength: Epistemic Confidence and Identity Centrality

*Neil Van Leeuwen**

Department of Philosophy, Neuroscience Institute, Georgia State University, Atlanta, GA, United States

Keywords: belief, credence, measurement, identity, epistemic confidence, identity centrality

INTRODUCTION: A DIFFERENCE WORTH MEASURING

What does it mean to have "strong beliefs"? My thesis is that it can mean two very different things. That is, there are two distinct psychological features to which "strong belief" can refer, and these often come apart. I call the first feature *epistemic confidence* and the second *identity centrality*. They are conceptually distinct and, if we take ethnographies of religion seriously, distinct in fact as well. If that's true, it's methodologically important for the psychological sciences to have *measures* that tease them apart.

EPISTEMIC CONFIDENCE VS. IDENTITY CENTRALITY

The following hypothetical case illustrates the distinction.

> Johan (a young Afrikaner man) frequently insists that his deceased father was an opponent of Apartheid in the 1970s and 1980s. This is part of the standard narrative of his family history he gives to people he meets. Yet privately he knows he doesn't have that much evidence it's true, and sometimes he suspects his father just told him that to make himself look good.
>
> Johan also has a great deal of certain knowledge of various things that matter little to him. He knows Istanbul used to be called Constantinople, one is technically not a prime, and Toyota has manufacturing facilities in South Africa.

In this case, identity centrality and epistemic confidence come apart in both directions. Johan has a *high* degree of identity centrality for the idea that his father opposed Apartheid (that idea is part of his social identity), but he has a *low* degree of epistemic confidence in that idea (he's not sure it's true). Conversely, he has a *high* degree of epistemic confidence that Istanbul used to be called Constantinople, but that same idea, for him, has a *low* degree of identity centrality—if it has any.

The difference between the two psychological features is also apparent in real-world ethnographies of religion. I focus here on ethnographies of The Vineyard Church (a large, representative American Evangelical sect[1]) by Luhrmann (2012) and Bialecki (2017). A central practice of the Vineyard is "hearing" the voice of God, where this typically involves having internal auditory mental imagery. Yet Vineyard members often speak of such experiences like this: "Sometimes when we think it's the spirit moving, it's just our burrito from lunch" and "There's always a choice to believe what it is" (Luhrmann, 2012, p. 70). Relatedly, Bialecki notes that Vineyard members commonly joke about the difficulty of determining whether the feelings they're having are from God or from the pizza they had for lunch. Hence, Vineyard members are often *unsure* (epistemically unconfident) that God spoke to them. Uncertainty is apparent also in that they describe it as a "choice" to believe and commonly find it *difficult* to "believe," as do members

[1]The Vineyard is now an international movement, but it originated in the US and still has many features characteristic of American Evangelicalism.

OPEN ACCESS

Edited by:
Rüdiger J. Seitz, Heinrich Heine University of Düsseldorf, Germany

Reviewed by:
Carlos Montemayor, San Francisco State University, United States

***Correspondence:**
Neil Van Leeuwen
nvan@gsu.edu

Specialty section:
This article was submitted to Theoretical and Philosophical Psychology, a section of the journal Frontiers in Psychology

Received: *10 May 2022*
Accepted: *30 May 2022*
Published: *29 June 2022*

Citation:
Van Leeuwen N (2022) Two Concepts of Belief Strength: Epistemic Confidence and Identity Centrality. Front. Psychol. 13:939949. doi: 10.3389/fpsyg.2022.939949

of other Christian sects (Appiah, 2019, p. 37–38). Vineyard members often struggle with *doubt*. One Vineyard member even said, "I don't believe it. But I'm sticking with it. That's my definition of faith" (Luhrmann, 2012, p. 316). I propose that this cluster of phenomena can best be explained by positing that many Vineyard members have a low degree of epistemic confidence in their "beliefs," while those "beliefs" are nevertheless central to their identities, which is why they keep showing up, putting in effort, and saying things like "I'm sticking with it." Without being confident that their "beliefs" describe how things really are, they maintain them because of who *they* are (cf. Heiphetz et al., 2014).

To capture the difference in question, let the following serve as working definitions that can be refined through iterative stages of empirical inquiry and theoretical reflection.

> Epistemic confidence: The degree to which someone feels a belief state approximates *knowledge*.

(Knowledge, for purposes of this definition, implies clear contents, objective truth, and rational justification[2].)

> Identity centrality: The degree to which someone experiences a belief state as part of their *social identity*.

(Social identity, for purposes of this definition, is a cluster of psychological states and behavioral dispositions that constitute someone as a member of an actual or potential in-group, or that an individual uses to achieve a desired social position[3].)

To be clear, I am not suggesting that identity centrality is more important than epistemic confidence, or *vice versa*. They are just different psychological features that should not be confused. So now let's examine how some current measures of "belief" in psychology of religion fare in light of this distinction[4].

On Fullerton and Hunsberger's (1982) "Christian Orthodoxy Scale," respondents write down integers ranging from −3 ("strongly disagree") to +3 ("strongly agree") next to various claims. For example: "God exists as: Father, Son, and Holy Spirit," "Man is *not* a special creature made in the image of God, he is simply a recent development in the process of animal evolution" (where this is *contrary* to Christian orthodoxy), "Jesus Christ was the divine Son of God," etc.

No doubt the Fullerton and Hunsberger scale captures something important about religious psychology in that it measures Christian orthodoxy in some sense. But suppose a researcher administered the scale to a group of participants and most of them put "+3" next to all orthodox items and "−3" next to all contra-orthodox items. Would that researcher know whether the "+3" and "−3" responses were driven more by *epistemic confidence* or more by *identity centrality*? She would not. The reason why is that *either* psychological feature could cause a participant to put down "+3" next to the orthodox items. People who are epistemically confident an idea is true will typically be motivated to express "strong agreement" with it (conversely for disagreement), but so will people for whom that idea is central to their identity. So the Fullerton and Hunsberger scale doesn't capture this important difference.

To put the point abstractly, for any proposition *p*, a person with a high degree of epistemic confidence that *p* and a person with a high degree of identity centrality for *p* are both likely to put "+3" next to a sentence expressing *p*. So the scale does not discriminate.

Furthermore, if we trust the ethnographies just mentioned, this is a domain in which we might expect the two features to come apart. A researcher might hypothesize that many orthodox Christians are high in identity centrality with respect to orthodox beliefs, while being low or lower in epistemic confidence. The scale itself, however, would not help *test* that hypothesis.

This is not the place for an exhaustive catalog of belief measures, but it is worth observing how some prominent measures tilt toward one psychological feature or another, while others are entirely ambiguous between them.

In developmental psychology, Paul Harris has initiated a cross-cultural research program that compares people's confidence (probed in various ways) in the existence of scientific entities (e.g., germs, oxygen, etc.) to their confidence in the existence of supernatural entities of their religions (God, angels, etc.). Findings indicate that, even in religious societies like the United States and Iran, children and adults alike generally have lower degrees of confidence in religious than in scientific entities (Harris et al., 2006; Davoodi et al., 2018; Clegg et al., 2019). Measurement instruments in this line of research tilt in the direction of tracking epistemic confidence, but it is hard to rule out that identity centrality is also playing a role in driving some of the "confident" responses concerning religious entities. Hence, the findings could *understate* people's difference in epistemic confidence concerning scientific and religious entities.

Within social psychology of religion, some measures do help track identity centrality. Lindeman et al. (2020), for example, have items that probe how desirable religiosity is for respondents, whether they take religion to be harmful, and the degree of strong emotions elicited by religion. They also ask directly: "How important are religious attitudes to your identity?" Such questions are indeed useful in tracking the identity centrality. But they do not offer much in terms of assessing whether and how epistemic confidence and identity centrality converge or diverge. They also do not assess those psychological features in relation to *specific* religious doctrines and stories, such as the existence of a triune God or the Virgin Birth.

In psychology of religion more generally, there are indeed measures that track adherence to specific belief contents, but (like Fullerton and Hunsberger) without distinguishing epistemic confidence and identity centrality. Jong et al. (2013), for example, include these items: "There exists an all-powerful, all-knowing,

[2]This characterization of "knowledge" is intended for purposes of this definition only and not as a fundamental analysis of that term. Note also that I am not saying that an epistemically confident belief state must *be* knowledge—only that the person who has it must *feel* (to some degree) like it approximates knowledge.

[3]The in-group in Johan's case would be socially liberal South Africans. I should also note that there are interesting relations between the two kinds of "strength." In particular, one uses one's perceived knowledge about one's group to figure out what "beliefs" should be accorded identity centrality. So epistemic confidence does play a foundational role in relation to identity centrality, even if identity centrality guides symbolic behaviors that often seem to go against it. This is similar to the anti-symmetric cognitive governance relation I discuss in my (2014) paper, and I thank one anonymous referee for raising the issue.

[4]I focus here on psychology of religion, but my points are likely to generalize to other sub-fields as well.

TABLE 1 | Dimensions of difference: epistemic confidence and identity centrality.

	Epistemic confidence	Identity centrality
Formation conditions	Constrained by cognition of evidence	Social opportunity (Stark and Finke, 2000), voluntary choice (Kierkegaard, 1843/1985; Luhrmann, 2012).
Extinction conditions	Cognition of contrary evidence	Value conflict with group leaders or group in general (Sauvayre, 2011; Bialecki, 2017).
Action output	Decision theoretic, instrumental	Symbolic, experiential (Van Leeuwen and Van Elk, 2019; Luhrmann, 2020), solidarity building (Sosis and Alcorta, 2003; Bulbulia, 2004, 2012; Alcorta and Sosis, 2005; Henrich, 2009).

loving God," "Some people will go to Heaven when they die," etc. Their scale runs from −4 (strongly disagree) to +4 (strongly agree). They write, "The two ends of the scale are therefore designed to indicate extreme disbelief or atheism... and confident belief... whereas the midpoint of the scale (i.e., 0) implies agnosticism or uncertainty" (496). Other measures of "belief" have similar Likert scales (Tobacyk, 2004; Pennycook et al., 2012; e.g., Lindeman et al., 2015). The "strongly" in these scales is what's problematic: strong in *which way*—epistemic confidence or identity centrality? We don't know.

DISCUSSION: WHERE DO WE GO FROM HERE?

But why, one might ask, should we *want* measures that separate those features? The answer is that the difference matters for both descriptive psychological research and for normative philosophical research.

With regards to descriptive psychology, I submit that epistemic confidence and identity centrality are likely to differ along the following practically important dimensions: formation conditions (how a given attitude is formed), extinction conditions (how a given attitude is extinguished), and action guidance (what sorts of behaviors that attitude generates and how). In other words, an epistemically confident belief that *p* is likely to be formed, be revised, and generate action differently from a belief that *p* that is central to one's identity.

The table above (**Table 1**) lays out likely differences. (Here, the properties of epistemic confidence are standard in literature on "degrees of belief" in decision theory and formal epistemology[5].)

Much more can be said about each of these dimensions of variation. But the broad outlines are clear: Epistemic confidence is, with various exceptions, likely to respond to evidence and guide instrumentally rational actions; identity centrality is likely to respond to social pressures and guide in-group-oriented behavior and self-presentation. So this is a distinction that makes a host of differences.

With regards to normative philosophical research, it is fair to say that this distinction raises a range of questions. The most basic one is this: Should the norms of evidence and truth that seem clearly to apply to epistemic confidence transfer over to identity centrality? In point of fact, it seems that identity centrality *is* far less constrained by evidence. But *should* it be? This is an important question, whose answer I don't know, that the present work at least puts us in a position to ask more clearly. And if it turns out that the proper norms for the respective psychological features do differ, it is even more important to develop measurement tools that would detect which of the two phenomena we are dealing with for any given "belief" set. Otherwise we wouldn't know which norms are applicable in any given case.

One of the reasons, I suspect, why such tools are lacking is that teasing out the distinction using survey instruments is likely to be extremely hard. That is why the difference is easier to notice in ethnographies, which incorporate observation of non-verbal behavior and of more nuanced verbal behavior.

Yet building measurement scales would still be worth the attempt, and I suggest that the above chart could be used to generate proxies for the features in question in relation to specific belief contents. To what extent is one's belief (say) that God is triune constrained by evidence vs. being voluntarily chosen? To what extent is it likely to be rejected due to contrary evidence vs. value conflict with group leaders? Does it guide instrumental or symbolic actions? No doubt any such scale would elicit some noise in addition to signal. Nevertheless, appropriate measurement tools could well be crafted that get at important and striking differences in the ways people can and do have "strong" beliefs.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

This article is funded by Dr. Rüdiger Seitz, *via* the Volkswagen Foundation, Siemens Healthineers, and the Betz Foundation.

ACKNOWLEDGMENTS

The author thanks his research assistants who helped with this project: David Casey, Bilal Khan, and Trevor Rukwava. For feedback on earlier drafts, the author thanks Franz Dietrich, Kenny Easwaran, Paul Harris, Jonathan Jong, Gordon Pennycook, Rüdiger Seitz, and a referee. The author thanks Bence Nanay and the audience at the Center for Philosophical Psychology for discussion of a version I presented there. The author thanks Rüdiger Seitz, Hans-Ferdinand Angel, and the other participants at the Credition 2021 conference for discussion of the very first version of this paper.

[5]See also Van Leeuwen, 2014, 2017a,b, 2018 for parallel comparisons between factual belief, which has epistemic confidence, and religious credence, which has identity centrality.

REFERENCES

Alcorta, C. S., and Sosis, R. (2005). Ritual, emotion, and sacred symbols. *Hum. Nat.* 16, 323–359. doi: 10.1007/s12110-005-1014-3

Appiah, K. A. (2019). *The Lies That Bind: Rethinking Identity*. New York, NY: W. W. Norton & Co.

Bialecki, J. (2017). *A Diagram for Fire*. Oakland, CA: University of California Press. doi: 10.1525/california/9780520294202.001.0001

Bulbulia, J. (2004). The cognitive and evolutionary psychology of religion. *Biol. Philos.* 19, 655–686. doi: 10.1007/s10539-005-5568-6

Bulbulia, J. (2012). Spreading order: religion, cooperative niche construction, and risky coordination problems. *Biol. Philos.* 27, 1–27. doi: 10.1007/s10539-011-9295-x

Clegg, J. M., Cui, Y. K., Harris, P. L., and Corriveau, K. H. (2019). God, germs, and evolution: Belief in unobservable religious and scientific entities in the U.S. and China. *Integr. Psychol. Behav. Sci.* 53, 93–106. doi: 10.1007/s12124-019-9471-0

Davoodi, T., Jamshidi-Sianaki, M., Abedi, F., Payir, A., Cui, Y. K., Harris, P. L., and Corriveau, K. H. (2018). Beliefs about religious and scientific entities among parents and children in Iran. *Soc. Psychol. Personal. Sci.* 10, 847–855. doi: 10.1177/1948550618806057

Fullerton, J. T., and Hunsberger, B. (1982). A unidimensional measure of Christian orthodoxy. *J. Sci. Study Relig.* 21, 317–326. doi: 10.2307/1385521

Harris, P. L., Pasquini, E. S., Duke, S., Asscher, J. J., and Pons, F. (2006). Germs and angels: the role of testimony in young children's ontology. *Dev. Sci.* 9, 76–96. doi: 10.1111/j.1467-7687.2005.00465.x

Heiphetz, L., Spelke, E. S., Harris, P. L., and Banaji, M. R. (2014). What do different beliefs tell us? An examination of factual, opinion-based, and religious beliefs. *Cognit. Dev.* 30, 15–29. doi: 10.1016/j.cogdev.2013.12.002

Henrich, J. (2009). The evolution of costly displays, cooperation and religion. *Evol. Hum. Behav.* 30, 244–260. doi: 10.1016/j.evolhumbehav.2009.03.005

Jong, J., Bluemke, M., and Halberstadt, J. (2013). Fear of death and supernatural beliefs: Developing a new Supernatural Belief Scale to test the relationship. *Eur. J. Pers.* 27, 495–506. doi: 10.1002/per.1898

Kierkegaard, S. (1843/1985). *Fear and Trembling. Tr. by Alastair Hannay*. London: Penguin Books.

Lindeman, M., Marin, P., Schjoedt, U., and van Elk, M. (2020). Nonreligious identity in three Western European countries: a closer look at nonbelievers' self-identifications and attitudes towards religion. *Int. J. Psychol. Relig.* 30, 288–303. doi: 10.1080/10508619.2020.1746984

Lindeman, M., Svedholm-Häkkinen, A. M., and Lipsanen, J. (2015). Ontological confusions but not mentalizing abilities predict religious belief, paranormal belief, and belief in supernatural purpose. *Cognition.* 134, 63–76. doi: 10.1016/j.cognition.2014.09.008

Luhrmann, T. M. (2012). *When God Talks Back: Understanding the American Evangelical Relationship With God*. New York, NY: Vintage Books. doi: 10.24260/alalbab.v1i1.16

Luhrmann, T. M. (2020). *How God Becomes Real: Kindling the Presence of Invisible Others*. Princeton, NJ: Princeton University Press. doi: 10.1515/9780691211985

Pennycook, G., Cheyne, J. A., Seli, P., Koehler, D. J., and Fugelsang, J. A. (2012). Analytic cognitive style predicts religious and paranormal belief. *Cognition.* 123, 335–346. doi: 10.1016/j.cognition.2012.03.003

Sauvayre, R. (2011). "Le changement de croyance extremes: du cadre cognitive aux conflits de valeurs," Revue européenne des sciences sociales. *Eur. J. Soc. Sci.* 49, 61–82. doi: 10.4000/ress.830

Sosis, R., and Alcorta, C. (2003). Signaling, solidarity, and the sacred: the evolution of religious behavior. *Evol. Anthropol.* 12, 264–274. doi: 10.1002/evan.10120

Stark, R., and Finke, R. (2000). *Acts of Faith: The Human Side of Religion*. Oakland, CA: University of California Press.

Tobacyk, J. J. (2004). A revised paranormal belief scale. *Int. J. Transpers. Stud.* 23, 94–98. doi: 10.24972/ijts.2004.23.1.94

Van Leeuwen, N. (2014). Religious credence is not factual belief. *Cognition.* 133, 698–715. doi: 10.1016/j.cognition.2014.08.015

Van Leeuwen, N. (2017a). Do religious "beliefs" respond to evidence? *Philos. Explor.* 20, 52–72. doi: 10.1080/13869795.2017.1287294

Van Leeuwen, N. (2017b). Two paradigms for religious representation: the physicist and the playground (a reply to Levy). *Cognition.* 164, 206–211. doi: 10.1016/j.cognition.2017.03.021

Van Leeuwen, N. (2018). The factual belief fallacy. *Contemp. Pragmatism.* 15, 319–343. doi: 10.1163/18758185-01503004

Van Leeuwen, N., and Van Elk, M. (2019). Seeking the supernatural: the interactive religious experience model. *Religion Brain Behav.* 9, 221–251. doi: 10.1080/2153599X.2018.1453529

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Van Leeuwen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Behavioral Neuroscience

OPINION
published: 28 June 2022
doi: 10.3389/fnbeh.2022.952736

Linking Agent Detection of Invisible Presences to the Self: Relevance for Religious and Spiritual Experiences

Neza Vehar[1], *Jevita Potheegadoo*[1] *and Olaf Blanke*[1,2]*

[1] *Laboratory of Cognitive Neuroscience, Center for Neuroprosthetics and Brain Mind Institute, Ecole Polytechnique Fédérale de Lausanne (EPFL), Geneva, Switzerland,* [2] *Department of Clinical Neurosciences, Geneva University Hospital, Geneva, Switzerland*

Keywords: sensorimotor processing, hallucinations, agent detection, neurology, robotics, anthropology, cognitive science of religion, cultural beliefs

INTRODUCTION

The intriguing experience that somebody is nearby when no one is actually present and cannot be seen or heard has been reported in many different contexts and has been referred to as the sense of presence, feeling of a presence, invisible presences, or presence hallucination (PH) (James, 1902; Critchley, 1979). PHs are often vivid experiences, have a clear location in space—with people frequently turning around to search for the invisible presence—and some even offering it a chair or food (Jaspers, 1913; Nightingale, 1982). PHs are a common theme in fiction, having been alluded to in the literature of divinity, occultism, and parapsychology (Green and McCreery, 1975; Critchley, 1979) and studied in history and anthropology (Solomonova et al., 2011; Wyatt et al., 2016). Following reports of PHs in extreme mountaineering (Smythe, 1935; Messner, 2003), long-distance solo-biking (Davie, 2013), solo-sailing (Suedfeld and Mocellin, 1987) and in shipwreck survivors (Critchley, 1943), PHs have also been investigated in psychology and medicine (Critchley, 1979; Brugger et al., 1996; Arzy et al., 2006). Initially described in psychiatry (Jaspers, 1913; Llorca et al., 2016), PHs have more recently been mostly investigated in neurological patients with epilepsy, stroke, neoplasia, and Parkinson's disease (PD) (Brugger et al., 1996; Fénelon et al., 2011).

However, despite its intriguing experiential characteristics and the broad academic and clinical interest, scientific studies and experimental data on PHs continue to be sparse. This is likely due to difficulties in investigating a spontaneously occurring phenomenon, the absence of experimental procedures able to induce PHs reliably in real time, and to their occurrence in the large majority of cases in situations not prone to empirical investigations (far from laboratories). Here we provide an overview of recent investigations in clinical neuroscience on PH and in neuroscience using methods to induce PH experimentally, linking them to altered self-monitoring and sensorimotor processing. We analyze selected spiritual-religious experiences associated with PH and propose a new extended account of PH, by integrating and extending the altered self-monitoring account with the prominent agent detection theory in spiritual-religious experiences (Guthrie, 1989; Barrett and Lanman, 2008). We conclude by proposing that the mechanism and the controlled induction of invisible presences will likely have an impact in clinical and fundamental neurosciences and may provide a powerful experimental approach in biological anthropology and the cognitive science of religion.

NEUROLOGY AND NEUROSCIENCE

The feeling of presence has long interested psychologists (James, 1902), psychiatrists (Jaspers, 1913), and neurologists (Critchley, 1979), and has recently also been investigated as a clinical symptom. PH has been reported to co-occur with temporoparietal tumors (Brugger et al., 1996),

OPEN ACCESS

Edited by:
Rüdiger J. Seitz, Heinrich Heine University of Düsseldorf, Germany

Reviewed by:
Andrew A. Fingelkurts, BM-Science, Finland

***Correspondence:**
Olaf Blanke
olaf.blanke@epfl.ch

Specialty section:
This article was submitted to Emotion Regulation and Processing, a section of the journal Frontiers in Behavioral Neuroscience

Received: *25 May 2022*
Accepted: *07 June 2022*
Published: *28 June 2022*

Citation:
Vehar N, Potheegadoo J and Blanke O (2022) Linking Agent Detection of Invisible Presences to the Self: Relevance for Religious and Spiritual Experiences. Front. Behav. Neurosci. 16:952736. doi: 10.3389/fnbeh.2022.952736

epilepsy (Critchley, 1979; Brugger et al., 1996), stroke (Blanke et al., 2014), or schizophrenia (Llorca et al., 2016; Stripeikyte et al., 2021). Lately, PH has been classified as a frequent early hallucination in Parkinson's disease (Fénelon et al., 2011; Bernasconi et al., 2021) and Lewy Body dementia (Nagahama et al., 2010). Early evidence about specific brain areas was reported by Arzy et al. (2006), where PH was induced by electrical stimulation in temporoparietal cortex. Interestingly, with repeated stimulations, the PH was perceived with varying attributes (unknown, identified) and with mental states of intentionality and perceived attempts of interference. These data were extended by work using lesion overlap analysis of neurological patients, highlighting the involvement of several cortical regions (Blanke et al., 2014). Due to the high frequency of PH in Parkinson's disease (i.e., Fénelon et al., 2000), a recent study used functional lesion network analysis in patients with Parkinson's disease (Bernasconi et al., 2021) and determined a fronto-temporal PH network, involving ventral premotor cortex, inferior frontal gyrus, and posterior superior temporal sulcus region.

Based on these data and a prominent model that describes hallucinations as a disturbance or misattribution of self-related predictive sensory signals (Fletcher and Frith, 2009), Blanke et al. (2014) developed a robotic stimulation system that exposed participants to temporally and spatially conflicting sensorimotor signals. Participants were asked to repeatedly move the front robot with their hand (motor, tactile, and proprioceptive signals) and received tactile feedback on their back (back robot), under conditions of sensory deprivation (**Figure 1**). Being subjected to such conflicting somatosensory-motor stimulations characterized by an additional delay between front and back robot (Bernasconi et al., 2021) elicits PH in healthy individuals (Blanke et al., 2014; Bernasconi et al., 2021; Dhanis et al., 2022). Such robot-induced PH (ri-PH) also allow to study whether certain functions are associated with PH. When ri-PH are elicited while participants carry out a second task, changes in several perceptual and cognitive functions occur, including auditory perception (Orepic et al., 2021), thought insertion (Serino et al., 2021), and cognitive processes (Faivre et al., 2020). Do neurological and ri-PH and the described somatosensory-motor mechanisms relate to presences and invisible agents in anthropology and the science of religion?

SPIRITUAL-RELIGIOUS EXPERIENCES AND PH

The occurrence of invisible or imperceptible supernatural presences is commonly positioned at the core of religious or spiritual belief systems (Luhrmann et al., 2021), with wide cross-cultural variability, ranging from angels, spirits, natural forces to gods. Such presences can be broadly distinguished into intended or voluntary presences (individual actively sought out PH) and unintended, spontaneous presences (individual did not seek to experience PH). Since experiencing supernatural presences is often judged as socially and personally desirable, it is actively sought-after *via* rituals (Otero, 2003; Johnson et al., 2015), ingestion of psychedelic compounds (Sayin, 2016), as well as kindled through training, prayer (Luhrmann and Morgain, 2012; Luhrmann, 2020), or deliberate interaction (Morton, 2020). For example, intended presences have been described by Luhrmann (2012, p. 148) in her anthropological field work with American Evangelicals, actively invoking God's presence ("to feel sensorily aware of God, as if God were a person who was physically present"). Presences with spiritual-religious connotations also happen involuntarily (clinical or non-clinical) and may have transformative effects on a person's life. Non-clinical supernatural presences may "visit-upon" an unsuspecting, healthy person in extreme social or environmental conditions (i.e., Messner, 2003) but may also occur during mundane situations. The latter caught the attention of James (1902), reporting people that "felt a consciousness of a presence in the room (…) not the consciousness of a live person, but of a spiritual presence" (p. 62). James explicitly links PH to the "religious sphere of experience, (where) many persons (…) possess the objects of their belief not in the form of mere conceptions which their intellect accepts as true, but rather in the form of quasi-sensible realities directly apprehended" (p. 64). In the clinical context, PH experiences with spiritual-religious aspects have been noted for a long time. Jaspers (1913) described a patient with schizophrenia who reported PH characterized by "the feeling that the soul of his deceased father is with him," that the fatherly presence "is behind him" (p. 153), interfering with the patient's life (akin to reports of experiencing presence of ancestral spirits or in the context of grief; Klass and Goss, 1999; Pérez, 2011). Another patient noted at the onset of an epileptic seizure that he felt, "overwhelmingly real," somebody standing by his side (not seen or heard) and that "God was about to take me home and that I had not to fear anything in the world" (Brugger et al., 1996, p. 116). Despite their different contexts, these reports indicate many PH similarities including spatial aspects (presence behind the person or shoulder, mirroring how angels or other spirits are often represented), psychological attributes of the presence such as strong familiarity, psychological affinity, and identification of presence, and specific intentions (leading a person somewhere; guiding in danger, interfering with a task). PH also mostly appears in low luminosity or contrast conditions, at night and in extreme or monotonous environments, devoid of sensory stimulation—like revelations on mountains (Arzy et al., 2005), Inuit igloo confinement to evoke spirits (Geiger, 2009) or in hermits who retreated to deserts in early Christianity (Suedfeld and Mocellin, 1987).

INVISIBLE PRESENCES, HYPERSENSITIVE AGENT DETECTION AND THE COGNITIVE SCIENCE OF RELIGION

In brief, Guthrie's anthropomorphism account (Guthrie, 1980, 1989, 2001) linked agent-detection to supernatural beliefs, arguing for a low-level perceptual tendency anthropomorphizing the environment and detecting the presence of humans in environments devoid of others. Generating such false-positive agent perceptions may be adaptive in human evolution, because

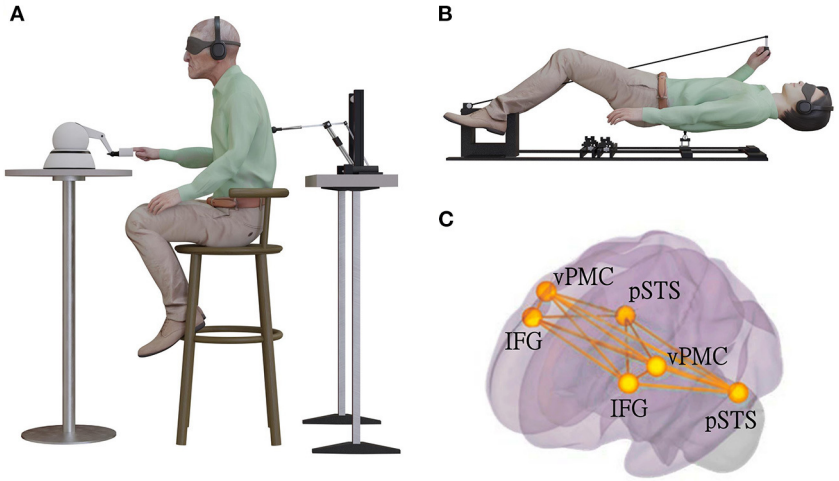


FIGURE 1 | Experimental induction of PH using sensorimotor stimulation (modified after Bernasconi et al., 2021). **(A)** The robotic set-up used for induction of invisible presences in patients with Parkinsons' disease in a sitting position. Patient was moving the front robot in a poking motion, using the right-hand index finger, while receiving a corresponding tactile stimulation on their back. They were in a state of sensory deprivation, wearing headphones playing white noise and with their eyes closed, wearing a blindfold. The stimulation on the back was either synchronous with their movement of the front robot (the back robot had 0 ms of delay) or asynchronous (the back robot randomly delayed from 0 to 500 ms in steps of 100 ms), with the asynchronous stimulation being significantly associated with experiencing robot-induced PH (ri-PH) as a function of sensorimotor delay. **(B)** The robotic set-up was adapted to be MR-compatible and used for an fMRI study in a healthy population of participants. **(C)** The brain activation and connectivity patterns were collected in an fMRI experiment with healthy and neurological non-parkinsonian patients. The schematic bilateral display of the connectivity overlap between the network connectivity in spontaneous PH identified using lesion network and connectivity analysis and ri-PH network from healthy participants. The bilateral regions are ventral premotor cortex (vPMC), inferior frontal gyrus (IFG), and posterior superior temporal sulcus (pSTS).

agents are sources of potential danger or opportunity (Van Leeuwen and van Elk, 2019). Accordingly, it has been argued that PH and supernatural agents (ghosts, gods, spirits) result from the recruitment of hyperactive perceptual mechanisms related to agent-detection. This account on over-detection of *humans* was extended to include *non-human agents* such as animals by Barrett (2000, 2011), and Barrett and Lanman (2008) and the broader detection of *agency* in the environment. As Guthrie's proposal did not account for the intentionality of presences, Barrett and colleagues proposed that agency-detection also involves perception of intentional states (motivations, intentions, desires), beyond mere detection of the presence, based on additional cognitive brain mechanisms (such as mentalizing).

Neuroscience data suggest a different account regarding invisibles presences: altered self-monitoring based on conflicting somatosensory-motor processes involving specific bodily signals. This self-monitoring approach is based on the misperception of oneself as another agent and was tested experimentally with ri-PH: The self is at the origin of invisible presences, being misperceived *as another agent*. We argue that invisible presences result not from visual-auditory mechanisms, as argued previously, but from a different perceptual mechanism: motor signals and their integration with somatosensory signals. These somatosensory-motor signals are specific and involve the global self-representation of a person's body (Blanke and Metzinger, 2009; Park and Blanke, 2019). The self-monitoring approach of PH sides with Guthrie that agent-detection is associated with perception, but primarily results from somatosensory-motor (not visual-auditory) perception, as shown in experiments applying somatosensory-motor stimulation in blindfolded noise-isolated participants (Blanke et al., 2014; Bernasconi et al., 2021; Dhanis et al., 2022; Orepic et al., 2021; Serino et al., 2021). Our proposal refines the anthropomorphic account and supports that the detection of *human agents* (self, global-body representation) and not the broader, less specific detection of *human and non-human agency* (animals, body-part representation) is key in PH. We strongly agree with Barrett that intention recognition, mentalizing, and the notion of minimally counterintuitive states are important to consider in PH. However, the involvement of (too) many different perceptual and cognitive functions conceptually seems to over-complexify matters, hindering empirical verification. Moreover, Barrett's proposed list of additional perceptual and cognitive mechanisms, may not be necessary to perceive invisible agents endowed with intentionality. As reviewed above, perceptual somatosensory-motor mechanisms related to a person's global self-representation are sufficient to perceive an intentional presence and fit well within the category of minimally counterintuitive states (Barrett and Lanman, 2008). Many fascinating questions remain. How does agent detection lead to religious beliefs, how is it shared among kin, and why do humans not simply discard these incorrect perceptions (Boyer, 2001; Van Leeuwen and van Elk, 2019)?

We conclude that the self-monitoring account of invisible presences is relevant not only in neurological and ri-PH, but also anthropology. It is perceptual in nature and links PH to *human agent* detection based on altered somatosensory-motor processing. It differs from previous accounts that have focused on altered perceptual or cognitive mechanisms related to the extrapersonal environment (unrelated to the observer's

somatosensory-motor body). Key aspects of supernatural agent detection, as noted by Barrett and Guthrie, are also accounted for by the self-monitoring proposal. We argue that these different accounts are not mutually exclusive: it is rather likely that self-related (egocentric) and environment-related (allocentric) mechanisms are complementary, although we argue that the self-monitoring account is the primary and most basic mechanism of supernatural agent-detection. Finally, the narrative style of both earlier theories has made it difficult to empirically test them, leading to numerous studies with conflicting results (Gervais et al., 2011; van Elk, 2013; van Elk et al., 2016; Maij et al., 2017, 2019). The new method of ri-PH (Blanke et al., 2014; Bernasconi et al., 2021; Serino et al., 2021; Dhanis et al., 2022) provides a promising way to investigate the role of the self in the intriguing human experience of supernatural agents, spirits, and gods. Future studies of invisible presences in different cultural contexts, integrating robot-induced PH with social science approaches, may facilitate the interaction of scholars from neuroscience, anthropology, and the cognitive science of religion.

AUTHOR CONTRIBUTIONS

NV and OB contributed to conception and theoretical basis for the article, and wrote sections of the manuscript. NV wrote the first draft of the paper and performed literature review. NV collected cases of PH experiences, under the guidance from OB. JP provided discussion and comments on the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

The authors are supported by two donors advised by CARIGEST SA (Fondazione Teofilo Rossi di Montelera e di Premuda and a third one wishing to remain anonymous), by National Center of Competence in Research (NCCR Synapsy 51NF40-185897), Parkinson Suisse, and Bertarelli Foundation. This paper is funded by Dr. Rüdiger Seitz, via the Volkswagen Foundation, Siemens Healthineers, and the Betz Foundation. The authors declare that Siemens Healthlineers was not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

ACKNOWLEDGMENTS

The authors would like to thank Arthur Trivier for the design of **Figure 1**, modified from Bernasconi et al. (2021).

REFERENCES

Arzy, S., Idel, M., Landis, T., and Blanke, O. (2005). Why revelations have occurred on mountains? Linking mystical experiences and cognitive neuroscience. *Med. Hypotheses* 65, 841–845. doi: 10.1016/j.mehy.2005.04.044

Arzy, S., Seeck, M., Ortigue, S., Spinelli, L., and Blanke, O. (2006). Induction of an illusory shadow person. *Nature* 443, 287. doi: 10.1038/443287a

Barrett, J. (2011). *Cognitive Science, Religion and Theology: From Human Minds to Divine Minds (Templeton Science and Religion Series)*. West Conshohocken, PA: Templeton Press.

Barrett, J. L. (2000). Exploring the natural foundations of religion. *Trends Cogn. Sci.* 4, 29–34. doi: 10.1016/S1364-6613(99)01419-9

Barrett, J. L., and Lanman, J. A. (2008). The science of religious beliefs. *Religion* 38, 109–124. doi: 10.1016/j.religion.2008.01.007

Bernasconi, F., Blondiaux, E., Potheegadoo, J., Stripeikyte, G., Pagonabarraga, J., Bejr-Kasem, H., et al. (2021). Robot-induced hallucinations in Parkinson's disease depend on altered sensorimotor processing in fronto-temporal network. *Sci. Transl. Med.* 13, 8362. doi: 10.1126/scitranslmed.abc8362

Blanke, O., and Metzinger, T. (2009). Full-body illusions and minimal phenomenal selfhood. *Trends Cogn. Sci.* 13, 7–13. doi: 10.1016/j.tics.2008.10.003

Blanke, O., Pozeg, P., Hara, M., Heydrich, L., Serino, A., Yamamoto, A., et al. (2014). Report neurological and robot-controlled induction of an apparition. *Curr. Biol.* 24, 2681–2686. doi: 10.1016/j.cub.2014.09.049

Boyer, P. (2001). *Religion Explained: The Evolutionary Origins of Religious Thought*. New York, NY: Basic Books.

Brugger, P., Regard, M., and Landis, T. (1996). Unilaterally felt "presences": the neuropsychiatry of one's invisible doppelgänger. *Neuropsychiatry Neuropsychol. Behav. Neurol.* 9, 114–122.

Critchley, H. (1943). *Shipwreck Survivors: A Medical Study*. London: J. & A. Churchill.

Critchley, M. (1979). *The Divine Banquet of the Brain and Other Essays*. New York, NY: Raven Press.

Davie, K. (2013). *Freedom Rider: 10,000km by Mountain Bike Across South Africa*. Johannesburg: Jacana Media, 244.

Dhanis, H., Blondiaux, E., Bolton, T., Faivre, N., Rognini, G., Van De Ville, D., et al. (2022). Robotically-induced hallucination triggers subtle changes in brain network transitions. *Neuroimage.* 248. doi: 10.1016/j.neuroimage.2021.118862

Faivre, N., Vuillaume, L., Bernasconi, F., Salomon, R., Blanke, O., and Cleeremans, A. (2020). Sensorimotor conflicts alter metacognitive and action monitoring. *Cortex* 124, 224–234. doi: 10.1016/j.cortex.2019.12.001

Fénelon, G., Mahieux, F., Huon, R., and Ziégler, M. (2000). Hallucinations in Parkinson's disease. Prevalence, phenomenology and risk factors. *Brain* 123, 733–745. doi: 10.1093/brain/123.4.733

Fénelon, G., Soulas, T., De Langavant, L. C., Trinkler, I., and Bachoud-Lévi, A. C. (2011). Feeling of presence in Parkinson's disease. *J. Neurol. Neurosurg. Psychiatry* 82, 1219–1224. doi: 10.1136/jnnp.2010.234799

Fletcher, P. C., and Frith, C. D. (2009). Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia. *Nat. Rev. Neurosci.* 10, 48–58. doi: 10.1038/nrn2536

Geiger, J. G. (2009). *The Third Man Factor: Surviving the Impossible, 2010th Edn*. Edinburgh: Canongate Books.

Gervais, W. M., Willard, A. K., Norenzayan, A., and Henrich, J. (2011). The cultural transmission of faith: why innate intuitions are necessary, but insufficient, to explain religious belief. *Religion* 41, 389–410. doi: 10.1080/0048721X.2011.604510

Green, C. E., and McCreery, C. (1975). *Apparitions*. London: Institute of Psychophysical Research.

Guthrie, S. (1980). A cognitive theory of religion. *Curr. Anthropol.* 21, 181–203. doi: 10.1086/202429

Guthrie, S. (2001). "Why gods? A cognitive theory," in *Religion in Mind: Cognitive Perspectives on Religious Belief, Ritual, and Experience*, ed J. Andersen (Cambridge University Press), 94–111. doi: 10.1017/CBO9780511586330.004

Guthrie, S. E. (1989). *Faces in the Clouds*. New York, NY: Oxford University Press.

James, W. (1902). *The Varieties of Religious Experience: A Study in Human Nature, 1st Edn*. New York, NY: Longmans, Green, and CO. Available online at: https://upload.wikimedia.org/wikipedia/commons/d/df/William_James_The_Varieties_of_Religious_Experience.pdf (accessed April 23, 2022).

Jaspers, K. (1913). Über leibhaftige Bewusstheiten (Bewusstheitstäuschungen), ein psychopathologisches Elementarsymptom. *Zeitschrift Pathopsychol.* 2, 150–161.

Johnson, K. A., Li, J., and Cohen, A. B. (2015). Fundamental social motives and the varieties of religious experience. *Relig. Brain Behav.* 5, 197–231. doi: 10.1080/2153599X.2014.918684

Klass, D., and Goss, R. (1999). Spiritual bonds to the dead in cross-cultural and historical perspective: comparative religion and modern grief. *Death Stud.* 23, 547–567. doi: 10.1080/074811899200885

Llorca, P. M., Pereira, B., Jardri, R., Chereau-Boudet, I., Brousse, G., Misdrahi, D., et al. (2016). Hallucinations in schizophrenia and Parkinson's disease: an analysis of sensory modalities involved and the repercussion on patients. *Sci. Rep.* 6, 1–9. doi: 10.1038/srep38152

Luhrmann, T. M. (2012). *When God Talks Back: Understanding the American Evangelical Relationship With God*. New York, NY: Vintage Books, division of Random House Inc. doi: 10.24260/alalbab.v1i1.16

Luhrmann, T. M. (2020). *How God Becomes Real: Kindling the Presence of Invisible Others*. Princeton, NJ: Princeton University Press. doi: 10.1515/9780691211985

Luhrmann, T. M., and Morgain, R. (2012). Prayer as inner sense cultivation: an attentional learning theory of spiritual experience. *Ethos* 40, 359. doi: 10.1111/j.1548-1352.2012.01266.x

Luhrmann, T. M., Weisman, K., Aulino, F., Brahinsky, J. D., Dulin, J. C., Dzokoto, V. A., et al. (2021). Sensing the presence of gods and spirits across cultures and faiths. *Proc. Natl. Acad. Sci. U. S. A.* 118, e2016649118. doi: 10.1073/pnas.2016649118

Maij, D. L. R., Van Harreveld, F., Gervais, W., Schrag, Y., Mohr, C., and Van Elk, M. (2017). Mentalizing skills do not differentiate believers from non-believers, but credibility enhancing displays do. *PLoS ONE* 12, e0182764. doi: 10.1371/journal.pone.0182764

Maij, D. L. R., van Schie, H. T., and van Elk, M. (2019). The boundary conditions of the hypersensitive agency detection device: an empirical investigation of agency detection in threatening situations. *Relig. Brain Behav.* 9, 23–51. doi: 10.1080/2153599X.2017.1362662

Messner, R. (2003). *The Naked Mountain*. Marlborough, UK: Crowood Press. Available online at: https://books.google.com/books/about/The_Naked_Mountain.html?id=HbdEPwAACAAJ (accessed May 13, 2022).

Morton, L. (2020). *Calling the Spirits: A history of Seances*. London: Reaktion Books.

Nagahama, Y., Okina, T., Suzuki, N., and Matsuda, M. (2010). Neural correlates of psychotic symptoms in dementia with Lewy bodies. *Brain* 133, 557–567. doi: 10.1093/brain/awp295

Nightingale, S. (1982). Somatoparaphrenia: a case report. *Cortex* 18, 463–467. doi: 10.1016/S0010-9452(82)80043-9

Orepic, P., Rognini, G., Kannape, O. A., Faivre, N., and Blanke, O. (2021). Sensorimotor conflicts induce somatic passivity and louden quiet voices in healthy listeners. *Schizophr. Res.* 231, 170–177. doi: 10.1016/j.schres.2021.03.014

Otero, R. A. (2003). The transformation of identity through possession rituals in popular religion. *Religion* 33, 249–262. doi: 10.1016/S0048-721X(03)00056-3

Park, H. D., and Blanke, O. (2019). Coupling inner and outer body for self-consciousness. *Trends Cogn. Sci.* 23, 377–388. doi: 10.1016/j.tics.2019.02.002

Pérez, E. (2011). Spiritist mediumship as historical mediation: African-American pasts, black ancestral presence, and Afro-Cuban religions. *J. Relig. Africa* 41, 330–365. doi: 10.1163/157006611X604760

Sayin, H. U. (2016). Psychoactive plants used during religious rituals. *Neuropathol. Drug Addict. Subst. Misuse* 3, 17–28. doi: 10.1016/B978-0-12-800634-4.00002-0

Serino, A., Pozeg, P., Bernasconi, F., Solcà, M., Hara, M., Progin, P., et al. (2021). Thought consciousness and source monitoring depend on robotically controlled sensorimotor conflicts and illusory states. *iScience* 24, 101955. doi: 10.1016/j.isci.2020.101955

Smythe, F. (1935). "The second assault," in *Attack on Everest*, ed H. Ruttledge (New York, NY: R.M. McBride), 227–228.

Solomonova, E., Frantova, E., and Nielsen, T. (2011). Felt presence: the uncanny encounters with the numinous Other. *AI Soc.* 26, 171–178. doi: 10.1007/s00146-010-0299-x

Stripeikyte, G., Potheegadoo, J., Progin, P., Rognini, G., Blondiaux, E., Salomon, R., et al. (2021). Fronto-temporal disconnection within the presence hallucination network in psychotic patients with passivity experiences. *Schizophr. Bull.* 47, 1718–1728. doi: 10.1093/schbul/sbab031

Suedfeld, P., and Mocellin, J. S. P. (1987). The "sensed presence" in unusual environments peter suedfeld. *Environ. Behav.* 19, 33–52. doi: 10.1177/0013916587191002

van Elk, M. (2013). Paranormal believers are more prone to illusory agency detection than skeptics. *Conscious. Cogn.* 22, 1041–1046. doi: 10.1016/j.concog.2013.07.004

van Elk, M., Rutjens, B. T., van der Pligt, J., and van Harreveld, F. (2016). Priming of supernatural agent concepts and agency detection. *Relig. Brain Behav.* 6, 4–33. doi: 10.1080/2153599X.2014.933444

Van Leeuwen, N., and van Elk, M. (2019). Seeking the supernatural: the Interactive Religious Experience Model. *Relig. Brain Behav.* 9, 221–251. doi: 10.1080/2153599X.2018.1453529

Wyatt, J., Tamas, S., and Bondi, L. (2016). Traces: an introduction to the special issue. *Emot. Sp. Soc.* 19, 37–39. doi: 10.1016/j.emospa.2016.02.005

Conflict of Interest: OB is an inventor on patent US 10,286,555 B2 (Title: Robot-controlled induction of the feeling of a presence) held by the Swiss Federal Institute (EPFL) that covers the robot-controlled induction of the feeling of a presence (PH). He is an inventor on patent US 10,349,899 B2 (Title: System and method for predicting hallucinations) held by the Swiss Federal Institute (EPFL) that covers a robotic system for the prediction of hallucinations for diagnostic and therapeutic purposes. He is a cofounder and shareholder of Metaphysiks Engineering SA, a company that develops immersive technologies, including applications of the robotic induction of PHs. He is member of the board and shareholder of Mindmaze SA.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Vehar, Potheegadoo and Blanke. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Behavioral Neuroscience

TYPE Opinion
PUBLISHED 06 October 2022
DOI 10.3389/fnbeh.2022.977600

Check for updates

OPEN ACCESS

EDITED BY
Rüdiger J. Seitz,
Heinrich Heine University of Düsseldorf, Germany

REVIEWED BY
Bruno J. Weder,
University of Bern, Switzerland

*CORRESPONDENCE
Irene Cristofori
icristofori@isc.cnrs.fr

SPECIALTY SECTION
This article was submitted to Emotion Regulation and Processing, a section of the journal Frontiers in Behavioral Neuroscience

RECEIVED 30 June 2022
ACCEPTED 20 September 2022
PUBLISHED 06 October 2022

CITATION
Cristofori I, Cohen-Zimerman S, Bulbulia J, Gordon B, Krueger F and Grafman J (2022) The neural underpinning of religious beliefs: Evidence from brain lesions. *Front. Behav. Neurosci.* 16:977600. doi: 10.3389/fnbeh.2022.977600

COPYRIGHT
© 2022 Cristofori, Cohen-Zimerman, Bulbulia, Gordon, Krueger and Grafman. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

The neural underpinning of religious beliefs: Evidence from brain lesions

Irene Cristofori[1,2]*, Shira Cohen-Zimerman[3,4], Joseph Bulbulia[5,6], Barry Gordon[7,8], Frank Krueger[9,10] and Jordan Grafman[3,4,11,12]

[1]Institute of Cognitive Sciences Marc Jeannerod CNRS, UMR 5229, Bron, France, [2]University of Lyon, Etablissement 1, Villeurbanne, France, [3]Cognitive Neuroscience Laboratory, Brain Injury Research, Shirley Ryan AbilityLab, Chicago, IL, United States, [4]Department of Physical Medicine and Rehabilitation, Northwestern University, Evanston, IL, United States, [5]School of Psychology, Faculty of Science, Victoria University of Wellington, Wellington, New Zealand, [6]Department of Linguistic and Cultural Evolution, Max Plank Institute for the Science of Human History, Jena, Germany, [7]Cognitive Neurology/Neuropsychology Division, Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, MD, United States, [8]Department of Cognitive Science, Johns Hopkins University, Baltimore, MD, United States, [9]School of Systems Biology, George Mason University, Fairfax, VA, United States, [10]Department of Psychology, George Mason University, Fairfax, VA, United States, [11]Departments of Neurology, Psychiatry, and Cognitive Neurology and Alzheimer's Disease, Feinberg School of Medicine, Chicago, IL, United States, [12]Department of Psychology, Northwestern University, Chicago, IL, United States

KEYWORDS

religious belief, cognition, social cognition, lesions mapping, dorsolateral prefrontal cortex, ventromedial prefrontal cortex

Religious beliefs emerged in conjunction with moral beliefs, political, and legal beliefs (Cristofori and Grafman, 2017). As pointed out by Oviedo and Szocik (Oviedo and Szocik, 2020) recent debates attempted to determine whether compared to other beliefs, religious beliefs have a specific (Van Leeuwen, 2014) or shared cognitive structure (Boudry and Coyne, 2016). Empirical evidence supports both positions (Levy, 2017; Van Leeuwen, 2018), while neuroscientific findings support a shared neural network between religious and other beliefs [e.g., political beliefs (Cristofori et al., 2015)]. Religious beliefs shape a person's character and influence daily decision-making and social interactions (Cristofori and Grafman, 2017). Religious beliefs are often concerned with the existence of supernatural agents and are often entangled with moral, political, economic, and legal beliefs that have collectively had a profound influence throughout known human history. Their pervasiveness and power have suggested they have a special status in the human brain. However, despite being much in need of investigation, studying the neurobiological basis of religious beliefs has proved difficult. Here, we focus on the contribution of a set of brain-lesions studies that shed light on the neurocognitive underpinnings of religious beliefs. We then compare the findings of these brain-lesions studies with results from neuroimaging studies.

The unique contribution of brain lesion studies

Lesion studies were among the first approaches to investigating how neural-anatomy relates to brain-function (Harlow, 1848). Since the 1990 s, however, functional neuroimaging studies have come to dominate functional-anatomic brain research

[e.g., (Fox and Raichle, 2007; Raichle, 2009)]. Despite their popularity, functional neuroimages studies have two limitations. First, while such studies can reveal correlations between brain activation and behavior (e.g., area X is active when behavior y is present), they are limited in their scope for providing causal understanding about the relationship (Siddiqi et al., 2021). The extent to which randomized controlled experiments may manipulate brain anatomy is clearly limited. Second, even where interventions may in principle yield causal inferential knowledge, in practice, most neuroimaging studies are underpowered [e.g., (Marek et al., 2022)]. On the other hand, the distribution of brain lesions across a population is typically a matter of chance. For this reason, by comparing religious cognition in a population with focal lesions and without lesions, causal identification if functional neuroanatomical relationships may be possible. Another virtue of functional neuroanatomical lesion studies of religious cognition, as we shall next describe, is that they are not underpowered. It is important to acknowledge that even whether brain lesions mapping is a model for studying structure/function relationship, lesions emerging from a traumatic brain injury can involve adjacent brain areas that, may accommodate different neurological and/or cognitive functions. For instance, see Rorden et al. recommendations (Rorden et al., 2009) for performing lesion behavior mapping. We do note, however, that we have studied *penetrating* traumatic brain injuries and, in that case, identification of the lesion parameters like volume loss and areas affected are made somewhat easier by the relative focality of the lesion (Raymont et al., 2011). In addition, we must highlight the studies presented here involve a slightly different cohort number, depending if the participant performed the test or not. As a reminder, our participants were tested over a 1-week period, with extensive neuropsychological and experimental assessments (for more details see Raymont et al., 2011).

The Vietnam head injury study and its contribution for understanding the neural basis of religiosity

The Vietnam Head Injury Study [VHIS (Raymont et al., 2011)] is a longitudinal follow-up of American male combat veterans who served in the Vietnam War, most of whom suffered from focal penetrating traumatic brain injury (pTBI). This dataset provide data on participants who are similar in age, and education level, and is unique in that it includes pre-injury intelligence. The study has followed the veterans (those with, and those without, focal brain lesions) for about 50 years post injury. Particularly, in the *final phase* of the VHIS (2008–2012), 169 participants (134 with pTBI, 35 with no injury) were assessed for executive functions, social cognition, personality, as well as large battery of tests dedicated to religious beliefs, including fundamentalism, God-Image, and mystical experiences. Individuals studied here are from a monotheistic culture with one God (or no God) rather than many Gods.

Highlighted findings

The ventromedial prefrontal cortex and its involvement in religious beliefs

Religious beliefs can reflect the particular relationship an individual has with God. A strong personal relationship with God is theoretically (Fiori et al., 2006) and empirically (Newton and McIntosh, 2010) associated with an enhanced sense of control. A recent study by Cohen-Zimerman et al. (2020) aimed at understanding whether damage to the vmPFC—a region associated with emotionally meaningful religious experiences and with a sense of control—could modulate self-reports of a personal relationship with God and a sense of control. Voxel-based lesion-symptom mapping found that damage to the right vmPFC caused a stronger personal relationship with God, and patients with damage to this region reported a greater sense of control compared to patients with damage to the posterior cortex as well as matched healthy patients. Moreover, the association between the vmPFC damage and a greater sense of control was associated with a stronger personal relationship with God. Taken together, these results suggest that a strong personal relationship with God can serve a crucial psychological function by affecting a sense of personal control, with both enhanced after right vmPFC lesions.

More recently, Cristofori et al. (2021) investigated the neural interplay between empathy and personal relationship with God. Extending previous observations that theory of mind networks is recruited during prayer (Boyer, 2003, 2008), the authors found that people with vmPFC damage reported higher scores on the personal relationship with God inventory (Lawrence, 1997) even when they were not praying. The results showed that vmPFC and posterior superior temporal sulcus/temporoparietal junction (pSTS/TPJ) lesions, associated with the strength of the personal relationship with God, affected empathetic responses. The authors suggested that the neurological networks underpinning God representations amplify human empathetic responses. The cultural evolutionary study of religion has argued that supernatural beliefs evoke pro-social responses because people fear the wrath of Gods (Atran and Norenzayan, 2004). In accordance with other studies e.g., (Norenzayan et al., 2012), our findings imply that, in contrast to the focus of the evolutionary literature on punishment, greater attention should be addressed to investigating the mechanisms by which the religious belief system modulates empathetic responses to others. It may seem that a *stronger* relationship with God, based on the lesioned right vmPFC, is counterintuitive. However, a stronger relationship with God post-injury might be due to the crucial role of the vmPFC in scaling and evaluating social behavior

(Moretti et al., 2009; Cristofori et al., 2015). Counterintuitive behavior changes provided by lesion-free contralateral homotopic areas are documented in the neuroscientific literature. However, recovery from severe and complex neural deficits may be more dependent upon extended neural networks rather than on a confined neural structure.

The dorsolateral prefrontal cortex and its involvement in religious beliefs

Religious beliefs can be influenced by certain experiences, such as mystical experiences, i.e., subjectively believed encounters with a supernatural world. Mystical experiences diverge from religious beliefs, in the sense that someone can experience a mystical phenomenon even without prior religious beliefs. Cristofori and collaborators (Cristofori et al., 2016) investigated pTBI patients and healthy volunteers. Mystical experiences were assessed using the Mystical scale [M-Scale (Hood, 1975)]. The M-scale refers to mystical experiences that the people may have previously experienced (e.g., "I have had an experience that was both timeless and spaceless"). Voxel-based lesion-symptom mapping analysis showed that lesions to frontal and temporal brain regions were linked with greater mystical experiences. Such regions included the dlPFC and middle/superior temporal cortex (TC). Performing a confirmatory group analysis, the researchers found that the dlPFC lesion group reported experiencing increased mysticism. Notably, longitudinal analysis of pre-injury data (correlating with general intelligence and executive functions task performance) excludes explanations from individual differences. These findings support previous speculation linking executive functions to mystical experiences and reveal that executive functions (particularly those aspects of executive functions that depend upon dlPFC) causally contribute to the down-regulation of mystical experiences. This study provided evidence in favor of the executive inhibition hypothesis, for the emergence of mystical experiences. This hypothesis was based on previous studies where the authors observed decreased activity in the dlPFC during mystical exercises in practitioners of glossolalia [i.e., religious prayer group experiences in which individuals speak an incomprehensible language (Newberg et al., 2006)] or a reduction of cognitive resources invested in error monitoring during religious rituals (Schjoedt et al., 2013). Another neuroimaging study has shown that participants down-regulated regions in the dlPFC during prayers performed by charismatic speakers (Schjoedt et al., 2011).

Among the different aspects that characterize religious beliefs, a crucial one is the strength of the beliefs, i.e., the fundamentalism aspect. Previous research has identified the vmPFC as critical to representing fundamentalism (Asp et al., 2012). However, the means by which vmPFC regulates fundamentalism was still less certain. Zhong and collaborators hypothesized that the vmPFC represents diverse religious beliefs and that a vmPFC lesion would be associated with religious fundamentalism or the narrowing of religious beliefs (Zhong et al., 2017). To test this prediction, the authors assessed religious adherence with a widely-used religious fundamentalism scale (Altemeyer and Hunsberger, 1992). The results showed that participants with dlPFC lesions had fundamentalism beliefs similar to patients with vmPFC lesions, however, the effect of a dlPFC lesion on fundamentalism was associated with decreased cognitive flexibility and openness. These findings indicated that cognitive flexibility and openness are necessary for flexible and adaptive religious commitment and that such diversity of religious thought is dependent on the functionality of the dlPFC.

The increase of fundamentalism and diminished flexibility and openness might be related to hemisphere dominance laterality. There is evidence that the left hemisphere is focused on facts whereas the right hemisphere represents contextual information and, therefore more adapted to specific *situations* (McGilchrist, 2012). Moreover, in mental tasks, like meditation, left-brain dominance is effective in top-down regulation, while interhemispheric integration of facts and context takes place in anterior brain areas (Raffone et al., 2019). In addition, other psychological factors such as emotional support and self-efficacy (Zahodne et al., 2014) may influence diminished cognitive resources and induce increased fundamentalism to control new experiences and daily life.

Conclusion

In summary, the lesion studies we describe above identify a network of neural substrates involved in the production of religious cognition, and clarify the functional relationship between religious beliefs, mystical experience, executive control,

TABLE 1 Represents a summary of how damage to a brain region affects religious beliefs/experience, cognition/social cognition.

Lesions	Religious belief/experience	Cognition/social cognition
vmPFC	↑ Relationship with god	↑ Control* & empathy
dlPFC	↑ Mystical experiences	↓ Executive functions
dlPFC	↑ Fundamentalism	↓ Flexibility & openness

Pointing up/down arrows indicated an increased/decreased religious or cognitive mechanism. *The relationship with God and sense of control involved specifically the right vmPFC.

emotional regulation, rigidity of ideological commitments, and other features of social cognition. The following Table 1 summarizes the main results.

We have delineated a network of lesioned areas within the prefrontal cortex including dlPFC and vmPFC, and more posterior regions such as superior temporal cortex (STC), and pSTS/TPJ (default mode network). More importantly, the network involved included areas damaged in both hemispheres, i.e., vmPFC on the right (belonging functionally to the default mode network) and dlPFC on both sides (belonging functionally to the fronto-parietal control network). In addition, spared areas of the left hemisphere might have driven recovery, e.g., left vmPFC homotopic to its damaged right counterpart.

To sum up, the studies reviewed here complement functional neuroimaging [e.g., (Schjoedt et al., 2011)] and non-invasive brain stimulation [e.g., (Holbrook et al., 2016)] in contributing to our understanding of the religious belief system.

Our results support the belief system model proposed by Seitz and collaborators (see Seitz et al., 2018). According to the *credition* model, beliefs are the result of neural processes involving the perception and evaluation of external information, and they drive individuals' decisions. Beliefs are unique representations with imaginative and emotional content, using linguistic and memory functions by which beliefs can be expressed, stored, and recalled (Seitz, 2022). Beliefs are fundamental cognitive constructs connecting people's prior experiences with their future behaviors (Krueger and Grafman, 2017).

Author's note

Questions concerning the Vietnam Head Injury Study can be directed to JG, jgrafman@northwestern.edu.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Funding

This research was supported by the Therapeutic Cognitive Neuroscience Fund (BG). The funder played no role in the design of this study or the interpretation of its results.

Acknowledgments

The authors thank all the Vietnam veterans, without their long-term commitment to improving the health care of veterans, these studies could not have been completed.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Altemeyer, B., and Hunsberger, B. (1992). Authoritarianism, religious fundamentalism, quest, and prejudice. *Int. J. Psychol. Relig.* 2, 113–133.

Asp, E., Ramchandran, K., and Tranel, D. (2012). Authoritarianism, religious fundamentalism, and the human prefrontal cortex. *Neuropsychology* 26, 414–421. doi: 10.1037/a0028526

Atran, S., and Norenzayan, A. (2004). Religion's evolutionary landscape: counterintuition, commitment, compassion, communion. *Behav. Brain Sci.* 27, 713–730. doi: 10.1017/S0140525X04000172

Boudry, M., and Coyne, J. (2016). Disbelief in belief: on the cognitive status of supernatural beliefs. *Philosoph. Psychol.* 29, 601–615. doi: 10.1080/09515089.2015.1110852

Boyer, P. (2003). Religious thought and behaviour as by-products of brain function. *Trends Cogn. Sci.* 7, 119–124. doi: 10.1016/s1364-6613(03)00031-7

Boyer, P. (2008). Religion: bound to believe? *Nature* 455, 1038–1039. doi: 10.1038/4551038a

Cohen-Zimerman, S., Cristofori, I., Zhong, W., Bulbulia, J., Krueger, F., Gordon, B., et al. (2020). Neural underpinning of a personal relationship with God and sense of control: A lesion-mapping study. *Cogn. Affect. Behav. Neurosci.* 20, 575–587. doi: 10.3758/s13415-020-00787-4

Cristofori, I., Bulbulia, J., Shaver, J. H., Wilson, M., Krueger, F., and Grafman, J. (2016). Neural correlates of mystical experience. *Neuropsychologia* 80, 212–220. doi: 10.1016/j.neuropsychologia.2015.11.021

Cristofori, I., and Grafman, J. (2017). "Neural underpinnings of the human belief system," in *Processes of Believing: The Acquisition, Maintenance, and Change in Creditions New Approaches to the Scientific Study of Religion*, eds. H.-F. Angel, L. Oviedo, R. F. Paloutzian, A. L. C. Runehov, and R. J. Seitz (Cham: Springer International Publishing), 111–123.

Cristofori, I., Viola, V., Chau, A., Zhong, W., Krueger, F., Zamboni, G., et al. (2015). The neural bases for devaluing radical political statements revealed by penetrating traumatic brain injury. *Soc. Cogn. Affect. Neurosci.* 10, 1038–1044. doi: 10.1093/scan/nsu155

Cristofori, I., Zhong, W., Cohen-Zimerman, S., Bulbulia, J., Gordon, B., Krueger, F., et al. (2021). Brain networks involved in the influence of religion on empathy in male Vietnam War veterans. *Sci. Rep.* 11:11047. doi: 10.1038/s41598-021-90481-3

Fiori, K. L., Brown, E. E., Cortina, K. S., and Antonucci, T. C. (2006). Locus of control as a mediator of the relationship between religiosity and life satisfaction: age, race, and gender differences. *Mental Health Rel. Culture* 9, 239–263. doi: 10.1080/13694670600615482

Fox, M. D., and Raichle, M. E. (2007). Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat. Rev. Neurosci.* 8, 700–711. doi: 10.1038/nrn2201

Harlow, J. M. (1848). Passage of an Iron rod through the head. *Boston Med. Surg. J.* 39, 389–393. doi: 10.1056/NEJM184812130392001

Holbrook, C., Izuma, K., Deblieck, C., Fessler, D. M. T., and Iacoboni, M. (2016). Neuromodulation of group prejudice and religious belief. *Soc. Cogn. Affect. Neurosci.* 11, 387–394. doi: 10.1093/scan/nsv107

Hood, R. W. (1975). The construction and preliminary validation of a measure of reported mystical experience. *J. Sci. Study Rel.* 14, 29–41. doi: 10.2307/1384454

Krueger, F., and Grafman, J. (2017). *The Neural Basis of Human Belief Systems*. Available online at: https://www.routledge.com/The-Neural-Basis-of-Human-Belief-Systems/Krueger-Grafman/p/book/9781138109896 (accessed August 28, 2022).

Lawrence, R. T. (1997). Measuring the image of god: the god image inventory and the god image scales. *J. Psychol. Theol.* 25, 214–226.

Levy, N. (2017). Religious beliefs are factual beliefs: content does not correlate with context sensitivity. *Cognition* 161, 109–116. doi: 10.1016/j.cognition.2017.01.012

Marek, S., Tervo-Clemmens, B., Calabro, F. J., Montez, D. F., Kay, B. P., Hatoum, A. S., et al. (2022). Reproducible brain-wide association studies require thousands of individuals. *Nature* 603, 654–660. doi: 10.1038/s41586-022-04492-9

McGilchrist, I. (2012). *The Master and His Emissary: The Divided Brain and the Making of the Western World*. Loondon: Yale University Press.

Moretti, L., Dragone, D., and di Pellegrino, G. (2009). Reward and social valuation deficits following ventromedial prefrontal damage. *J. Cogn. Neurosci.* 21, 128–140. doi: 10.1162/jocn.2009.21011

Newberg, A. B., Wintering, N. A., Morgan, D., and Waldman, M. R. (2006). The measurement of regional cerebral blood flow during glossolalia: a preliminary SPECT study. *Psychiatry Res. Neuroimag.* 148, 67–71. doi: 10.1016/j.pscychresns.2006.07.001

Newton, A. T., and McIntosh, D. N. (2010). Specific religious beliefs in a cognitive appraisal model of stress and coping. *Int. J. Psychol. Rel.* 20, 39–58. doi: 10.1080/10508610903418129

Norenzayan, A., Gervais, W. M., and Trzesniewski, K. H. (2012). Mentalizing deficits constrain belief in a personal god. *PLoS ONE* 7, e36880. doi: 10.1371/journal.pone.0036880

Oviedo, L., and Szocik, K. (2020). Religious—and other beliefs: how much specificity? *Sage Open* 10:2158244019898849. doi: 10.1177/2158244019898849

Raffone, A., Marzetti, L., Del Gratta, C., Perrucci, M. G., Romani, G. L., and Pizzella, V. (2019). Toward a brain theory of meditation. *Prog. Brain Res.* 244, 207–232. doi: 10.1016/bs.pbr.2018.10.028

Raichle, M. E. (2009). A brief history of human brain mapping. *Trends Neurosci.* 32, 118–126. doi: 10.1016/j.tins.2008.11.001

Raymont, V., Salazar, A. M., Krueger, F., and Grafman, J. (2011). "Studying injured minds" – the vietnam head injury study and 40 years of brain injury research. *Front. Neurol.* 2:15. doi: 10.3389/fneur.2011.00015

Rorden, C., Fridriksson, J., and Karnath, H-O. (2009). An evaluation of traditional and novel tools for lesion behavior mapping. *Neuroimage* 44, 1355–1362. doi: 10.1016/j.neuroimage.2008.09.031

Schjoedt, U., Sørensena, J., Nielboa, K. L., Xygalatasa, D., Mitkidisa, M., and Bulbulia, J. (2013). Cognitive resource depletion in religious interactions. *Religion Brain Behav.* 3, 39–86. doi: 10.1080/2153599X.2012.736714

Schjoedt, U., Stodkilde-Jorgensen, H., Geertz, A. W., Lund, T. E., and Roepstorff, A. (2011). The power of charisma–perceived charisma inhibits the frontal executive network of believers in intercessory prayer. *Soc. Cogn. Affect. Neurosci.* 6, 119–127. doi: 10.1093/scan/nsq023

Seitz, R. J. (2022). Beliefs: a challenge in neuropsychological disorders. *J Neuropsychol* 16, 21–37. doi: 10.1111/jnp.12249

Seitz, R. J., Paloutzian, R. F., and Angel, H.-F. (2018). From believing to belief: a general theoretical model. *J. Cogn. Neurosci.* 30, 1254–1264. doi: 10.1162/jocn_a_01292

Siddiqi, S. H., Schaper, F. L. W. V. J., Horn, A., Hsu, J., Padmanabhan, J. L., Brodtmann, A., et al. (2021). Brain stimulation and brain lesions converge on common causal circuits in neuropsychiatric disease. *Nat. Human Behav.* 5, 1707–1716. doi: 10.1038/s41562-021-01161-1

Van Leeuwen, N. (2014). Religious credence is not factual belief. *Cognition* 133, 698–715. doi: 10.1016/j.cognition.2014.08.015

Van Leeuwen, N. (2018). The factual belief fallacy. *Contem. Prag.* 15, 319–343. doi: 10.1163/18758185-01503004

Zahodne, L. B., Nowinski, C. J., Gershon, R. C., and Manly, J. J. (2014). Which psychosocial factors best predict cognitive performance in older adults? *J. Int. Neuropsychol. Soc.* 20, 487–495. doi: 10.1017/S1355617714000186

Zhong, W., Cristofori, I., Bulbulia, J., Krueger, F., and Grafman, J. (2017). Biological and cognitive underpinnings of religious fundamentalism. *Neuropsychologia* 100, 18–25. doi: 10.1016/j.neuropsychologia.2017.04.009

frontiers | Frontiers in Psychology

TYPE Brief Research Report
PUBLISHED 06 December 2022
DOI 10.3389/fpsyg.2022.935278

Check for updates

OPEN ACCESS

EDITED BY
Rüdiger J. Seitz,
Heinrich Heine University of Düsseldorf, Germany

REVIEWED BY
Matthias Bock,
Hospital "F. Tappeiner", Italy
Wojciech Malchrzak,
Wroclaw Medical University, Poland

*CORRESPONDENCE
Nina Dalkner
nina.dalkner@medunigraz.at

†These authors have contributed equally to this work and share first authorship

SPECIALTY SECTION
This article was submitted to Theoretical and Philosophical Psychology, a section of the journal Frontiers in Psychology

RECEIVED 03 May 2022
ACCEPTED 01 August 2022
PUBLISHED 06 December 2022

CITATION
Dalkner N, Fleischmann E, Fellendorf FT, Wagner-Skacel J, Schönthaler EMD, Bengesser S, Häussl A, Tietz S, Tmava-Berisha A, Lenger M and Reininghaus EZ (2022) COVID-19 vaccination motivation and underlying believing processes: A comparison study between individuals with affective disorder and healthy controls.
Front. Psychol. 13:935278.
doi: 10.3389/fpsyg.2022.935278

COPYRIGHT
© 2022 Dalkner, Fleischmann, Fellendorf, Wagner-Skacel, Schönthaler, Bengesser, Häussl, Tietz, Tmava-Berisha, Lenger and Reininghaus. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

COVID-19 vaccination motivation and underlying believing processes: A comparison study between individuals with affective disorder and healthy controls

Nina Dalkner[1*†], Eva Fleischmann[1†], Frederike T. Fellendorf[1], Jolana Wagner-Skacel[2], Elena M. D. Schönthaler[1], Susanne Bengesser[1], Alfred Häussl[1], Sophie Tietz[1], Adelina Tmava-Berisha[1], Melanie Lenger[1] and Eva Z. Reininghaus[1]

[1]Clinical Department of Psychiatry and Psychotherapeutic Medicine, Medical University Graz, Graz, Austria, [2]Clinical Department of Medical Psychology, Psychosomatic, and Psychotherapy, Medical University Graz, Graz, Austria

Background: Believing processes represent fundamental brain functions between cognition and emotion. Shortly before the introduction of a compulsory vaccination against COVID-19 in Austria, motives and underlying believing processes regarding the vaccination were collected in individuals with affective disorder (AD) and healthy controls (HC).

Methods: 79 individuals with AD and 173 HC were surveyed online to assess believing processes with the parameters of the credition model (narratives, certainty, emotion, mightiness) about (1) the coronavirus itself and (2) why someone is vaccinated or not. In addition, we calculated congruence scores between content of narrative and type of emotion and divided the narrative content into positive, negative, and indifferent.

Results: There were no differences in vaccination status between AD and HC. Higher levels of certainty were observed in HC compared to AD in both vaccinated and unvaccinated individuals. The effects were higher when asked about the motivation to vaccinate or not than about the coronavirus itself. In HC, more positive emotions and more congruence between emotions and narratives were reported during believing in their vaccination motives. No group differences were found in mightiness for both items. Independently from diagnosis, unvaccinated people had high levels of certainty and more negative emotions and narratives while believing in their motives for not getting vaccinated.

Conclusion: When believing about the COVID-19 vaccination, individuals with AD were more uncertain and experienced fewer positive emotions than HC, although both groups did not differ in vaccination status. These effects were not that strong when believing about the coronavirus in general.

KEYWORDS

COVID-19 vaccination, affective disorder, cognition, emotion, credition

Introduction

In Austria, containment measures against the coronavirus disease (COVID)-19 issued by the government included the obligation to be vaccinated or recovered when in public. In the period from December 1, 2021, to January 31, 2022, the number of individuals tested positive for COVID-19 increased from 1,175,785 to 1,891,468, and the number of deaths from or with COVID-19 increased from 12,458 to 13,669 (AGES, 2022; Epidemiologisches Meldesystem, 2022b). Starting on November 15, 2021, a lockdown for unvaccinated individuals was introduced, which lasted until January 31, 2022 (Niederösterreichische Nachrichten, 2022). A general lockdown was imposed from November 22 to December 11, 2021. On January 11, 2022, the decision of a nationwide vaccination obligation was proclaimed and with February 5, 2022, compulsory COVID-19 vaccination was required for adults aged 18 and older (Bundesministerium für Soziales, Gesundheit, Pflege und Konsumentenschutz, 2022a), which has been suspended again since March 9, 2022 (Bundesministerium für Soziales, Gesundheit, Pflege und Konsumentenschutz, 2022b). Up until January 31, 2022, 75.9% of Austrians had been vaccinated at least once, 72.1% had been vaccinated twice, and 49.8% had received the third shot (Epidemiologisches Meldesystem, 2022a). In comparison, 63.3% of Europeans had been vaccinated two times (Our World in Data, 2022).

Vaccination rates of individuals with psychiatric disorder were lower than those of the general population, despite having been given priority status in some countries (Tzur Bitan et al., 2021; Arumuham et al., 2022; Curtis et al., 2022). One reason for this might be vaccine hesitancy, which was more pronounced in individuals with mental illness than in healthy controls (HC; Hao et al., 2021; Jefsen et al., 2021; Eyllon et al., 2022). Factors associated with vaccine hesitancy were misinformation, fear (Payberah et al., 2022; Peritogiannis et al., 2022), mistrust (Payberah et al., 2022), and negative attitudes towards vaccines (Danenberg et al., 2021). Believing in the safety of vaccines and a good preventive effect were associated with vaccination willingness in individuals with psychiatric disorder (Huang et al., 2021).

Believing is a cognitive process consisting of formation, revision, and evaluation of beliefs (Angel and Seitz, 2016; Connors and Halligan, 2017). Credition describes the dynamic process of believing (Angel, 2013; Paloutzian and Mukai, 2017) as an interface between cognition and emotion. The credition model by Angel and Seitz (2016) encompasses four major parameters: proposition, certainty, emotion, and mightiness. The content of the statement about a certain belief is called "proposition." A person's inclination to believe the proposition is referred to as "certainty." The affective valence of the proposition is termed "emotion." The degree of significance of the proposition is termed "mightiness."

In a recent study during the COVID-19 pandemic, our study group demonstrated that credition parameters highly differed between patients with bipolar disorder and HC (Tietz et al., 2022). As the attitude towards the COVID-19 vaccination and motives to get vaccinated of individuals with psychiatric disorder remain largely unexplored and the underlying cognitive processes are unknown, we aimed to investigate believing processes around COVID-19 vaccination and to compare patients with affective disorders (AD) and HC. Additionally, we aimed to test for differences in believing depending on vaccination status (vaccinated or not vaccinated), as the understanding of believing processes (narrative, certainty, emotion, and mightiness) can provide a better overview of the motivators for vaccination and consequently increase the vaccination rate of people who are particularly at risk.

Materials and methods

An online survey was conducted with LimeSurvey (GmbH, 2003) and a link was sent out *via* e-mail to a pool of currently and previously treated patients at the Department of Psychiatry and Psychotherapeutic Medicine in Graz and was also shared *via* social media. The survey took place from December 14, 2021 to January 31, 2022. The study was approved by the local ethics committee and informed consent was given prior to study participation. In sum, 356 people opened the survey (104 of them indicated having a psychiatric disorder), and 252 (79 AD and 173 HC) of them filled out all items and were included in the analyses. The participants were surveyed on their vaccination status, demographic data, and with two questions in German language concerning their individual beliefs. The items of interest are listed in Table 1.

In addition to the proposition (narrative), the degree of certainty, the experienced emotion while believing (evaluated *via* an Emotion Wheel, see Figure 1), and the mightiness (strength of emotion) were assessed. As certainty and mightiness were rating scales, emotion was categorized into positive (happy), negative (sad, angry, anxious, disgusted), and indifferent (surprised) emotions. In addition, it was evaluated whether the narrative was positive, negative, or indifferent, and whether it matched the emotion (congruent) or not (incongruent).

Although the survey was sent out to former patients of the Department of Psychiatry and Psychotherapeutic Medicine, some diagnoses of AD (unipolar or bipolar affective disorder) were self-reported, as the link was additionally shared *via* social media for volunteers (Facebook and WhatsApp). HC had to state no psychiatric disorder themselves or in first-degree relatives (see control items).

Statistics

A multivariate analysis of co-variance (MANCOVA) with group (AD vs. HC) as independent variable controlling for age was calculated to test for between-subject differences in the

TABLE 1 Items of interest (Believing processes, vaccination status, psychiatric diagnosis control items).

I. Believing processes

Item 1 COVID-19 beliefs:

a. Proposition: *When I think about the coronavirus (COVID-19), I believe that … (narrative)*

b. Certainty: *On a scale from 0 (not sure) to 100 (quite sure), how sure are you about that while believing?*

c. Emotion: *Using the Emotion Wheel, please identify an emotion that most closely relates to your state while you are believing:…*

d. Mightiness: *On a scale of 0 (not at all) to 10 (very much), how strongly do you experience the emotion while believing?*

Item 2 Vaccination/Non-vaccination motive beliefs:

a. Proposition: *I am vaccinated/not vaccinated against COVID-19, because I believe that … (narrative)*

b. Certainty: *On a scale from 0 (not sure) to 100 (quite sure), how sure are you about you about that while believing?*

c. Emotion: *Using the Emotion Wheel, please identify an emotion that most closely relates to your state while you are believing:…*

d.Mightiness: *On a scale of 0 (not at all) to 10 (very much), how strongly do you experience the emotion while believing?*

II. Vaccination status

Have you been vaccinated against COVID-19? [Yes, fully immunized (at least 2 vaccinations)/Yes, one vaccination/No]

III. Psychiatric Diagnosis Control Items

1. *Please indicate which psychiatric disorder(s) you currently have (multiple answers possible): [None/Depressive disorder/Bipolar disorder/Panic disorder/Generalized Anxiety Disorder/Schizophrenia/Eating disorder/Alcohol use disorder/Other substance use disorder/Personality disorders/Other]*

2. *Please indicate which psychiatric disorder(s) you have ever been diagnosed with (multiple answers possible): [None/Depressive disorder/Bipolar disorder/Panic disorder/Generalized Anxiety Disorder/Schizophrenia/Eating disorder/Alcohol use disorder/Other substance use disorder/Personality disorders/Other]*

3. *Do you have first-degree relatives with a severe mental disorder (schizophrenia, bipolar disorder, major depressive disorder) [Yes/No]*

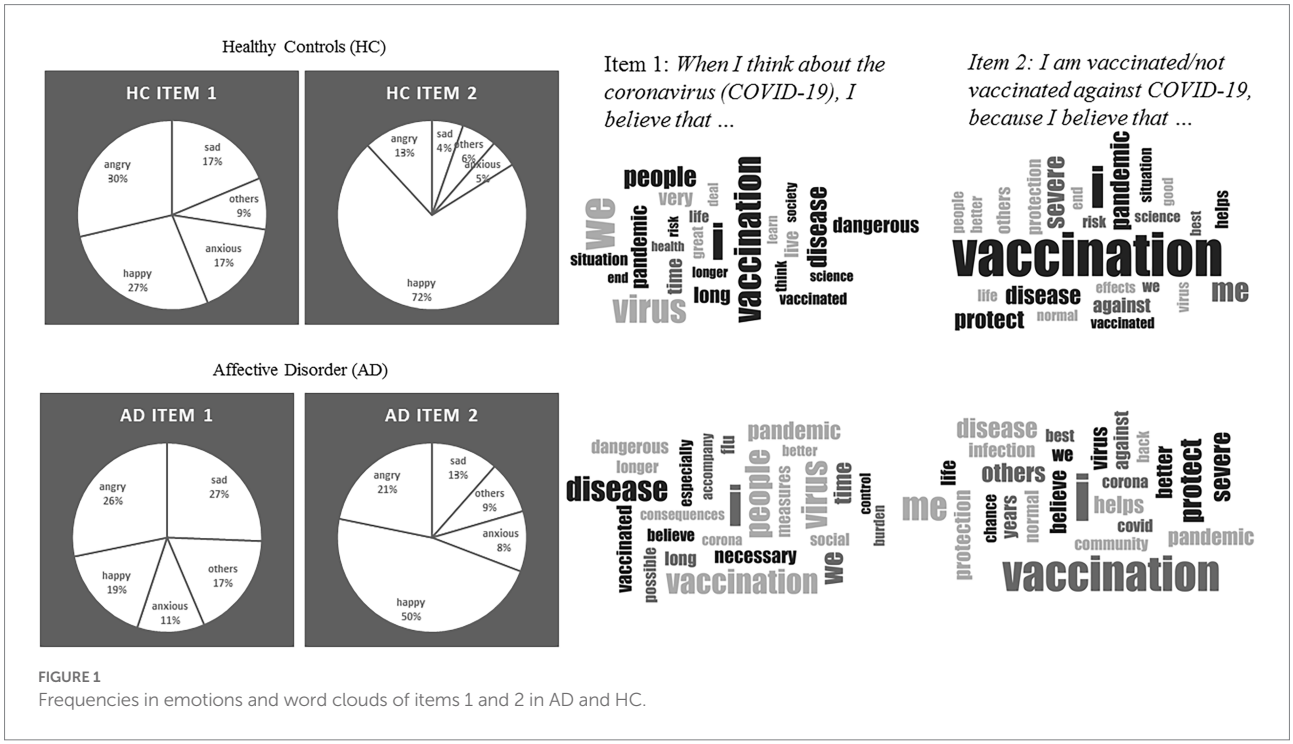


FIGURE 1
Frequencies in emotions and word clouds of items 1 and 2 in AD and HC.

credition parameters certainty and mightiness for both credition items. The *a priori* power analyses with Gpower 3.1.9.7. revealed for MANOVA (Global effects) a total sample size of $n = 252$, given effect size 0.0625, Power = 0.95, and Alpha = 0.05. As cell distribution was unequal in vaccination status, we could not enter this variable as second factor into the model but used *t*-tests to test for differences between vaccinated vs. non-vaccinated individuals (as homogeneity of variance was given). Differences in congruence (between emotion and narrative yes vs. no), emotion (positive = happy; negative = anxious, disgusted, angry, sad; indifferent = surprised), and narrative (positive, negative, indifferent) were calculated with chi-square tests and two-tailed Fisher's exact tests when more than 20% of expected frequencies were > 5. MANCOVA assumptions (normal distribution, homogeneity of variance) were checked. In ANOVA models, partial eta square (η^2_p), and for *t*-tests, Cohen's d as measure of effect size are presented. The obtained data were analyzed using IBM SPSS Statistics for Windows version 26.0 (Armonk,

New York: IBM Corp). In addition, we created word clouds in MAXQDA 2020 (VERBI Software, 2019) to present propositions for each item in the groups. Prepositions and conjunctions were ignored and added to a stop list in MAXQDA. The word clouds were translated from German into English for this study.

Results

In AD, the mean age was 43.78 years, 66.3% were female, and the median years of education were 16.20 years (see Table 2). In HC, the mean age was 37.17 years, 67.3% were female, and the median years of education were 16.7 years. Patients were significantly older than controls.

In AD, 96.2%, and in HC, 90.4% were vaccinated, i.e., had received at least one vaccination. Regarding immunization, i.e., having received at least two vaccinations, 3.8% of patients with AD and 10.2% of HC were not vaccinated or had only received one vaccination.

Differences between AD and HC

In response to COVID-19 beliefs in general (item 1), there was no multivariate effect [$F(2,248) = 1.24$, $p = 0.291$, $\eta^2_p = 0.01$; see Table 3]. Chi-square tests showed that HC had more positive emotions and fewer indifferent emotions than individuals with AD, who showed more indifferent emotions. There was no difference between the groups in frequencies of congruence or content of narrative.

Regarding item 2 Vaccination/Non-vaccination motive beliefs there was a significant multivariate group effect [$F(2,248) = 5.19$, $p = 0.006$, $\eta^2_p = 0.04$; see Table 3] indicating higher certainty in HC than in AD. No group effects were shown in mightiness. In addition, emotion differed between AD and HC, the latter reporting more positive and less negative emotions (see Figure 2). Furthermore, there was more congruence between emotion and narrative in HC than in AD. No group differences were shown in content of narratives.

Figure 3 shows the frequencies in emotions and the word clouds in AD vs. HC.

Differences between vaccinated and non-vaccinated individuals

T-tests showed significantly higher certainty levels in non-vaccinated people (item 1: $M = 90.6$, $SD = 15.7$) than in vaccinated people [$M = 83.9$, $SD = 16.4$; $t(255) = 1.75$, $p = 0.041$, Cohen's d = 0.35]. This effect was slightly higher for item 2 Vaccination/Non-vaccination motive beliefs [Non-vaccinated individuals: $M = 94.1$, $SD = 9.9$ vs. vaccinated individuals: $M = 89.1$, $SD = 14.8$; $t(254) = 2.04$, $p = 0.026$, Cohen's d = 0.39]. No group differences in mightiness were observed item 1: $t(255) = 0.05$, $p = 0.482$; item 2: $t(254) = 0.33$, $p = 0.370$.

Vaccinated people showed highly more positive emotions (68.2%) when believing about their motives for vaccination in comparison to non-vaccinated individuals, who reported more negative emotions while believing in their motives for non-vaccination [63.2%; $\chi^2(2) = 13.60$, $p = 0.001$]. No group differences were found in emotion in item 1 [$\chi^2(2) = 2.02$, $p = 0.364$].

In addition, non-vaccinated people showed highly more negative (85% vs. 4.6%) and fewer positive narratives (10.0% vs. 93.7%) than vaccinated individuals [Fisher's exact test: $\chi^2(1) = 125.70$, $p < 0.001$] for item 2. No group differences were found for item 1 [$\chi^2(2) = 3.80$, $p = 0.149$].

Two-tailed Fisher's exact tests showed that frequencies of congruence did not differ between vaccinated and non-vaccinated

TABLE 2 Sociodemographic characteristics of individuals with affective disorder and healthy controls.

Variables	Group		Test statistic (t, χ^2)	*p*-value	Cohen's *d*
	AD ($n = 79$)	HC ($n = 173$)			
Age ($M \pm SD$)	43.44 (13.86)	37.24 (13.36)	$t(249) = -3.38$	**<0.001**	−0.46
Sex (n, %)			$\chi^2(1) = 0.05$	0.818	
Female	52 (65.8%)	118 (67.3%)			
Male	27 (34.2%)	55 (31.8%)			
Median years of education ($M \pm SD$)	16.25 (6.24)	16.66 (3.60)	$t(102,469) = 0.55$	0.586	0.09
Vaccination status (n, %)			$\chi^2(1) = 2.75$	0.098	
Vaccinated[a]	77 (97.5%)	157 (90.8%)			
Unvaccinated	2 (2.5%)	16 (9.2%)			
Immunization against COVID-19 (n, %)			$\chi^2(1) = 1.94$	0.164	
Immunized[b]	76 (96.2%)	156 (90.2%)			
Not immunized	3 (3.8%)	17 (9.8%)			

AD = Affective disorder, HC = Healthy controls.
[a]At least one shot.
[b]At least two shots.
Bold value indicates a statistically significant difference $p<0.05$ between AD and HC.

TABLE 3 Descriptive statistics of the believing parameters of individuals with affective disorders and healthy controls.

Variables	Group		Test statistic	*p*-value	η_p^2
	AD (*n* = 79)	HC (*n* = 173)			
COVID-19 pandemic in general [a]					
Narratives (*n*, %)			$\chi^2(2) = 5.17$	0.075	
Positive	20 (25.3%)	57 (32.9%)			
Negative	57 (72.2%)	102 (59.0%)			
Indifferent	2 (2.5%)	14 (8.1%)			
Emotions (*n*, %)			$\chi^2(2) = 8.78$	**0.012**	
Positive	14 (17.7%)	47 (27.2%)			
Negative	52 (65.8%)	116 (67.1%)			
Indifferent	13 (16.5%)	10 (5.8%)			
Congruence[b] (*n*, %)			$\chi^2(1) = 0.43$	0.440	
Congruent	56 (70.9%)	131 (75.7%)			
Incongruent	23 (29.1%)	42 (24.3%)			
Certainty[d] (*M* ± *S*)	83.30 (15.32)	85.80 (16.64)	$F(1,249) = 2.11$	0.147	0.01
Mightiness[d] (*M* ± *SD*)	66.04 (26.26)	69.47 (23.80)	$F(1,249) = 0.93$	0.337	0.00
Vaccination [e]					
Narratives (*n*, %)			$\chi^2(1) = 0.77$[c]	0.771	
Positive	70 (88.6%)	152 (87.9%)			
Negative	8 (10.1%)	20 (11.6%)			
Indifferent	1 (1.3%)	1 (0.6%)			
Emotions (*n*, %)			$\chi^2(2) = 9.97$	**0.007**	
Positive	40 (50.6%)	123 (71.1%)			
Negative	34 (43.0%)	43 (24.9%)			
Indifferent	5 (6.3%)	7 (4.0%)			
Congruence[b] (*n*, %)			$\chi^2(1) = 3.84$	0.071	
Congruent	53 (67.1%)	136 (78.6%)			
Incongruent	26 (32.9%)	37 (21.4%)			
Certainty[c] (*M* ± *SD*)	85.58 (17.08)	91.79 (12.83)	$F(1,249) = 10.38$	**0.001**	0.04
Mightiness[c] (*M* ± *SD*)	74.67 (23.94)	77.62 (19.76)	$F(1,249) = 1.43$	0.233	0.01

AD = Affective disorder, HC = Healthy controls.
[a]When I think about the coronavirus (COVID-19), I believe that…
[b]Congruence between the narratives and the emotions.
[c]Fisher's exact test was used.
[d]In percent.
[e]I am vaccinated/not vaccinated against COVID-19, because I believe that….
Bold value indicates a statistically significant difference p<0.05 between AD and HC.

people [item 1: $\chi^2(1) = 1.31$, $p = 0.631$; item 2: $\chi^2(1) = 4.09$, $p = 0.165$].

Discussion

At the same time as the decision to introduce mandatory COVID-19 vaccinations in Austria in December 2021/January 2022, we surveyed 79 individuals with AD and 173 mentally healthy people. Their attitudes and beliefs about the coronavirus and their motives for vaccination/vs. non-vaccination were assessed using the parameters of the credition model (Angel, 2013; Angel and Seitz, 2016).

Individuals with AD and HC did not differ in their vaccination status. This has been shown in former Austrian studies with other samples (Fellendorf et al., 2022) as well as in international studies (Batty et al., 2021; Hao et al., 2021; Jefsen et al., 2021), although other studies showed a lower vaccine rate in individuals with psychiatric disorders (Arumuham et al., 2022; Curtis et al., 2022). We suppose a strong influence of socioeconomic circumstances, e.g., age, sex, education, and income, as well as cultural factors, such as governmental regulations in vaccination decision (Schwarzinger et al., 2021; Schernhammer et al., 2022). For example, both groups did not differ in education, although individuals with AD generally have lower levels of education (Lorant et al., 2003), which is more often found in unvaccinated individuals (Troiano and Nardi, 2021). There were also no differences in terms of sex. In this case, it would have been important to consider women's lower vaccine uptake (Troiano and Nardi, 2021). In relation to government regulations, the lockdown for unvaccinated people as well as the upcoming obligatory vaccination could have strongly encouraged both Austrian HC

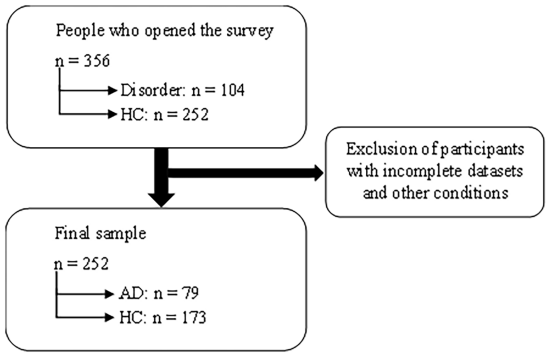


FIGURE 2
Flow chart of the participant selection.

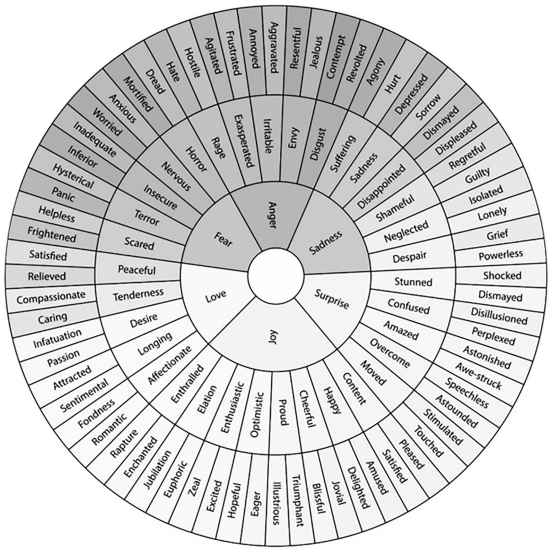


FIGURE 3
Emotion wheel used to select the emotion while believing.

and individuals with AD to get vaccinated. Moreover, there are no or only minimal private costs for healthcare in Austria for the individual, and although there were supply shortages, an easier general access to healthcare than in other countries could have further contributed to the results.

The COVID-19 pandemic is a highly emotional topic that is very much polarizing (Alam et al., 2021; Liew and Lee, 2021). This was also supported by the present study's results. When thinking about the coronavirus, HC reported more positive and less indifferent emotions while believing than individuals with AD. This is consistent with other studies that found that individuals with psychiatric disorders experienced more distress during the pandemic than HC (Solé et al., 2021). However, two thirds of both individuals with AD and HC reported negative emotions (anger, sadness, anxiety) when believing about the coronavirus, highlighting the continued negative influence of the pandemic on the population even at the beginning of 2022.

When thinking about their motives of vaccination, individuals with AD reported more negative emotions while believing than HC, most of whom reported positive emotions. Comparably, other studies found less vaccine acceptance in individuals with mental illness (Danenberg et al., 2021; Huang et al., 2021; Payberah et al., 2022), which is linked to negative feelings about the vaccination (de Vries et al., 2022). However, as vaccination rate did not differ in this study, emotions supposedly might not have played the essential role for individuals with AD when deciding whether they wanted to get vaccinated.

The results further showed that individuals with AD were less certain about their beliefs, especially regarding the COVID-19 vaccination. We assume that patients with AD have developed greater insecurity about potential threats based on their existing chronic mental disease, which could also lead to more self-care or a more ambivalent/incongruent attitude according to the stress-vulnerability model. Other possible reasons for our results might be mistrust, misinformation, and heightened fear, which has been shown to relate to vaccination hesitancy in individuals with mental illness (Payberah et al., 2022; Peritogiannis et al., 2022). The finding that individuals with a psychiatric disorder show less certainty about what they believe has also been observed in our first credition study in a sample of bipolar disorder (Tietz et al., 2022).

Independently from diagnosis, lower levels of certainty were also observed in vaccinated compared to non-vaccinated individuals. We assume that someone who is not vaccinated decides so with greater conviction (than someone who is vaccinated), and very strong negative emotions go along with it as supported by our findings. This goes in line with results by de Vries et al. (2022) demonstrating that individuals with vaccine hesitancy were less convinced.

of the emotional and rational advantages of COVID-19 vaccination and expressed more negative feelings about it. However, underlying reasons for non-vaccinations, including beliefs, have to be explored in samples with larger sample sizes.

This study has the following limitations. One problem of online studies is the sampling bias, such that only data from individuals who were motivated to participate in the survey were collected. This explains why most participants were vaccinated at least once and the group of unvaccinated was rather small. As vaccination rate in Austria was 70% at this time, there was a higher likelihood to recruit vaccinated people in a random sample (Epidemiologisches Meldesystem, 2022a). Thus, the cell sizes between vaccinated and unvaccinated individuals were too small to perform further statistical calculations, e.g., a 2×2 design with group and vaccination status would have been desirable. In addition, the diagnoses of AD were self-reported, but several control items were included. Moreover, instead of believing processes themselves, only verbal expressions could be examined. Believing processes might have been influenced by the subjects' introspective ability, which was not measured in the study. Furthermore, qualitative data had to be reduced by transforming

into positive, negative, and indifferent. It should also be noted that information may have been lost because of translation.

In conclusion, people with AD were more uncertain and experienced fewer positive emotions when thinking about their beliefs in the COVID-19 vaccination than HC. However, as both groups did not differ in vaccination rate, sociopolitical circumstances were presumably more influential in the decision to get vaccinated. Unvaccinated people were more likely to display negative emotions and narratives accompanied by high levels of certainty while believing in their motives for not getting vaccinated, but not when believing in the coronavirus in general; however, the cases of unvaccinated individuals were too small to draw final conclusions.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by Medical University of Graz, Austria. The patients/participants provided their written informed consent to participate in this study.

Author contributions

ND designed the study. ND and EF performed literature research as well as data analysis and wrote the first draft. FF, JW-S, ES, SB, AH, ST, AT-B, ML, and ER were responsible for proof reading and revising the manuscript. ES additionally supported the implementation of the study *via* the online application tool LimeSurvey. ER supervised the study procedure and revised important intellectual content. All authors contributed to the article and approved the submitted version.

Funding

The authors declare that this study received funding from the Volkswagen Foundation and the Betz Foundation. The authors declare that this study received funding from Siemens Healthineers. The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article, or the decision to submit it for publication.

Acknowledgments

We thank Lena Hiendl for her help in data analysis and Hans-Ferdinand Angel and the editor for their valuable suggestions on the content of this work.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

AGES (2022). AGES dashboard COVID19. Available at: https://covid19-dashboard.ages.at/ (Accessed March 13, 2022).

Alam, K. N., Khan, M. S., Dhruba, A. R., Khan, M. M., Amri, J. F., and Masud, M. (2021). Deep learning-based sentiment analysis of COVID-19 vaccination responses from twitter data. *Comput. Math. Methods Med.* 2021:4321131. doi: 10.1155/2021/4321131

Angel, H.-F. (2013). Credition. In: Anne L. C. Runehov and Lluis Oviedo (Eds.): *Encyclop. Sci. Rel.*, 2013, 536–539. doi: 10.1007/978-1-4020-8265-8_1565

Angel, H.-F., and Seitz, R. J. (2016). Process of believing as fundamental brain function: the concept of Credition. *SFU Res. Bulletin* 3, 1–20. doi: 10.15135/2016.4.1.1-20

Arumuham, A., O'Brien, Z., Ahmad, Z., Nikbin, K., and Howes, O. D. (2022). Low COVID-19 vaccination rates in people with severe mental illness and reasons for this: An out-patient study. *Acta Psychiatr. Scand.* 145, 416–418. doi: 10.1111/acps.1340

Batty, G. D., Deary, I. J., and Altschul, D. (2021). Pre-pandemic mental and physical health as predictors of COVID-19 vaccine hesitancy: evidence from a UK-wide cohort study. medRxiv 54, 274–282. doi: 10.1101/2021.04.27.21256185

Bundesministerium für Soziales, Gesundheit, Pflege und Konsumentenschutz (2022a). *Die Impfstoffe*. Available at: https://www.sozialministerium.at/Corona-Schutzimpfung/Corona-Schutzimpfung---Haeufig-gestellte-Fragen/Corona-Schutzimpfung---Haeufig-gestellte-Fragen---Die-Impfstoffe.html (Accessed March 13, 2022).

Bundesministerium für Soziales, Gesundheit, Pflege und Konsumentenschutz (2022b). *Impfpflicht*. Available at: https://www.sozialministerium.at/Corona-Schutzimpfung/Impfpflicht.html (Accessed March 13, 2022).

Connors, M. H., and Halligan, P. W. (2017). "Belief and belief formation: insights from delusions," in *Processes of Believing: The Acquisition, Maintenance, and change in Creditions*. eds. H.-F. Angel, L. Oviedo, R. F. Paloutzian, A. L. C. Runehov and R. J. Seitz (Basel: Springer International Publishing AG), 153–165.

Curtis, H. J., Inglesby, P., Morton, C. E., Mackenna, B., Green, A., and Hulme, W. (2022). Trends and clinical characteristics of COVID-19 vaccine recipients. *Br. J. Gen. Pract.* 1, 51–62. doi: 10.1101/2021.01.25.21250356

Danenberg, R., Shemesh, S., Tzur, D., Maoz, H., and Saker, T. (2021). Attitudes of patients with severe mental illness towards COVID-19 vaccinations: A preliminary report from a public psychiatric hospital Renana. *J. Psychiatr. Res.* 143, 16–20. doi: 10.1016/j.jpsychires.2021.08.020

de Vries, H., Verputten, W., Preissner, C., and Kok, G. (2022). COVID-19 vaccine hesitancy: the role of information sources and beliefs in Dutch adults. *Int. J. Environ. Res. Public Health* 19:3205. doi: 10.3390/ijerph19063205

Epidemiologisches Meldesystem (2022a). *COVID-19: Zeitreihe der verabreichten Impfdosen der Corona-Schutzimpfung*. Available at: https://www.data.gv.at/katalog/dataset/276ffd1e-efdd-42e2-b6c9-04fb5fa2b7ea

Epidemiologisches Meldesystem (2022b). *Katalog COVID-19: Zeitverlauf der gemeldeten COVID-19 Fälle im EMS (Morgenmeldung)*. Available at: https://www.data.gv.at/katalog/dataset/9723b0c6-48f4-418a-b301-e717b6d98c92 (Accessed March 12, 2022).

Eyllon, M., Dang, A. P., Barnes, J. B., Buresh, J., Peloquin, G. D., Hogan, A. C., et al. (2022). Associations between psychiatric morbidity and COVID-19 vaccine hesitancy: An analysis of electronic health records and patient survey. *Psychiatry Res.* 307:114329. doi: 10.1016/j.psychres.2021.114329

Fellendorf, F. T., Bonkat, N., Platzer, M., Schönthaler, E., Ratzenhofer, M., Bengesser, S. A., et al. (2022). Willingness to be vaccinated against COVID-19 is equal in individuals with affective disorders and healthy controls. *Vaccine:X* 11:100186. doi: 10.1016/j.jvacx.2022.100186

GmbH, LimeSurvey. (2003). *LimeSurvey*. Hamburg: LimeSurvey GmbH.

Hao, F., Wang, B., Tan, W., Husain, S. F., McIntyre, R. S., Tang, X., et al. (2021). Attitudes toward COVID-19 vaccination and willingness to pay: comparison of people with and without mental disorders in China. *BJPsych Open*. 7:e146. doi: 10.1192/bjo.2021.979

Huang, H., Zhu, X. M., Liang, P. W., Fang, Z. M., Luo, W., Ma, Y. M., et al. (2021). COVID-19 vaccine uptake, acceptance, and hesitancy among persons with mental disorders during the second stage of China's nationwide vaccine rollout. *Frontiers Med.* 8:761601. doi: 10.3389/fmed.2021.761601

Jefsen, O. H., Kølbæk, P., Gil, Y., Speed, M., Dinesen, P. T., Sønderskov, K. M., et al. (2021). COVID-19 vaccine willingness amongst patients with mental illness compared with the general population. *Acta neuropsychiatr.* 33, 273–276. doi: 10.1017/neu.2021.15

Liew, T. M., and Lee, C. S. (2021). Examining the utility of social media in COVID-19 vaccination: unsupervised learning of 672, 133 twitter posts. *JMIR Public Health Surveill.* 7:e29789. doi: 10.2196/29789

Lorant, V., Deliège, D., Eaton, W., Robert, A., Philippot, P., and Ansseau, M. (2003). Socioeconomic inequalities in depression: a meta-analysis. *Am. J. Epidemiol.* 157, 98–112. doi: 10.1093/aje/kwf182

Niederösterreichische Nachrichten (2022). Lockdown für Ungeimpfte endet am Montag. Available at: https://www.noen.at/in-ausland/lockdown-fuer-ungeimpfte-endet-am-montag-human-interest-innenpolitik-konjunktur-wien-oesterreich-310732109 (Accessed March 12, 2022).

Our World in Data (2022). Coronavirus (COVID-19) Vaccinations. Available at: https://ourworldindata.org/covid-vaccinations (Accessed March 13, 2022).

Paloutzian, R. F., and Mukai, K. (2017). "Believing, remembering, and imaging: The roots and fruits of meanings made and remade," in *Processes of Believing: The Acquisition, Maintenance, and change in Creditions*. eds. H.-F. Angel, L. Oviedo, R. F. Paloutzian, R, A. L. C. Runehov and R. J. Seitz (Basel: Springer International Publishing AG), 39–49.

Payberah, E., Payberah, D., Sarangi, A., and Gude, J. (2022). COVID-19 vaccine hesitancy in patients with mental illness: strategies to overcome barriers—a review. *J. Egypt. Public Health Assoc.* 97:5. doi: 10.1186/s42506-022-00102-8

Peritogiannis, V., Drakatos, I., Gioti, P., and Garbi, A. (2022). Vaccination rates against COVID-19 in patients with severe mental illness attending community mental health services in rural Greece. *Int. J. Soc. Psychiatry* 002076402210818. doi: 10.1177/00207640221081801

Schernhammer, E., Weitzer, J., Laubichler, M. D., Birmann, B. M., Bertau, M., Zenk, L., et al. (2022). Correlates of COVID-19 vaccine hesitancy in Austria: trust and the government. *J. Public Health (Oxf).* 44, e106–e116. doi: 10.1093/pubmed/fdab122

Schwarzinger, M., Watson, V., Arwidson, P., Alla, F., and Luchini, S. (2021). COVID-19 vaccine hesitancy in a representative working-age population in France: a survey experiment based on vaccine characteristics. *Lancet Public Health* 6, e210–e221. doi: 10.1016/S2468-2667(21)00012-8

Solé, B., Verdolini, N., Amoretti, S., Montejo, L., Rosa, A. R., Hogg, B., et al. (2021). Effects of the COVID-19 pandemic and lockdown in Spain: comparison between community controls and patients with a psychiatric disorder. Preliminary results from the BRIS-MHC STUDY. *J. Affect. Disord.* 281, 13–23. doi: 10.1016/j.jad.2020.11.099

Tietz, S., Wagner-Skacel, J., Angel, H.-F., Ratzenhofer, M., Fellendorf, F. T., Fleischmann, E., et al. (2022). Believing processes during the COVID-19 pandemic in individuals with bipolar disorder: An exploratory study. *World J. Psychiatry* 12, 929–943. doi: 10.5498/wjp.v12.i7.929

Troiano, G., and Nardi, A. (2021). Vaccine hesitancy in the era of COVID-19. *Public Health* 194, 245–251. doi: 10.1016/j.puhe.2021.02.025

Tzur Bitan, D., Kridin, K., Cohen, A. D., and Weinstein, O. (2021). COVID-19 hospitalisation, mortality, vaccination, and postvaccination trends among people with schizophrenia in Israel: a longitudinal cohort study. *Lancet Psychiatry* 8, 901–908. doi: 10.1016/S2215-0366(21)00256-X

VERBI Software (2019). *MAXQDA 2020*. Berlin: VERBI Software.

frontiers | Frontiers in Psychology

TYPE Opinion
PUBLISHED 07 October 2022
DOI 10.3389/fpsyg.2022.1010012

Check for updates

OPEN ACCESS

EDITED BY
Rüdiger J. Seitz,
Heinrich Heine University of Düsseldorf, Germany

REVIEWED BY
Eva Z. Reininghaus,
Medical University of Graz, Austria

*CORRESPONDENCE
Leonhard Schilbach
leonhard.schilbach@lvr.de

SPECIALTY SECTION
This article was submitted to Cognition, a section of the journal Frontiers in Psychology

RECEIVED 02 August 2022
ACCEPTED 29 August 2022
PUBLISHED 07 October 2022

CITATION
Pott J and Schilbach L (2022) Tracking and changing beliefs during social interaction: Where computational psychiatry meets cognitive behavioral therapy. *Front. Psychol.* 13:1010012. doi: 10.3389/fpsyg.2022.1010012

COPYRIGHT
© 2022 Pott and Schilbach. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Tracking and changing beliefs during social interaction: Where computational psychiatry meets cognitive behavioral therapy

Jennifer Pott[1] and Leonhard Schilbach[2,3]*

[1]Institut für Klinische Verhaltenstherapie, LVR-Klinikum Düsseldorf, Düsseldorf, Germany, [2]Abteilung für Allgemeine Psychiatrie 2, LVR-Klinikum Düsseldorf, Düsseldorf, Germany, [3]Medizinische Fakultät, Ludwig-Maximilians-Universität München, Munich, Germany

KEYWORDS

beliefs, computational psychiatry, cognitive behavioral therapies, psychosis, mechanisms

Introduction

Models of dysfunctional beliefs and belief systems have a long tradition in psychiatry and psychotherapy and have been used to explain and help treat various psychiatric disorders. In this article, we focus on the role of beliefs and belief updating in psychotic disorders, but also discuss how these phenomena extend into the normal and various other patient populations. In addition, we review insights from the field of "computational psychiatry," an area of research that uses mathematical models to describe and mechanistically explain how beliefs are formed, maintained or changed over time. We close by describing how cognitive behavioral therapy (CBT) uses the notion of beliefs to help treat psychiatric disorders and how an integration with "computational psychiatry" and digital phenotyping may help to provide new perspectives.

How beliefs can help to explain psychiatric disorders

Having beliefs about oneself and states of the world is indispensable for human life, because they allow us to constrain behavior even when we are faced with incomplete sensory information about the environment (Seitz, 2022). Here, beliefs can be defined as relatively stable accounts of what a subject holds to be true and anticipates to happen in the future, even though this typically takes the form of a probabilistic representation, because we can be more or less sure about something. These probabilistic representations are typically formed below awareness, but they powerfully influence emotions and actions in often predictable or sometimes even inflexible ways. Also, people tend to trust their beliefs and may even do so in the presence of conflicting evidence (Fletcher and Frith, 2009). In extreme cases, persons may even hold "fixed beliefs that are not amenable to change in light of conflicting evidence." Such beliefs according to the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-V) are characteristic of what in psychiatric terms would be described as a delusion. Delusions, in turn, are a common

feature of schizophrenia and other so-called psychotic disorders that can cause a person to lose touch with reality.

Research has investigated whether these so-called positive symptoms in schizophrenia are related to abnormal perception and/or abnormal beliefs. Both appear to be relevant, but according to a seminal review by Fletcher and Frith (2009) can be traced back to the same underlying core abnormality, i.e, a disturbance in error-dependent updating of inferences and beliefs about the world, which can be conceptualized as a disturbed hierarchical Bayesian framework. In such a framework–as introduced above–a belief is the subjective probability that some proposition about the world is true. This probability is continually updated in light of new incoming sensory evidence. Abnormal belief formation occurs when beliefs are not updated appropriately on the basis of new evidence (Hemsley and Garety, 1986). In line with these ideas, it has repeatedly been shown that persons with a diagnosis of schizophrenia show the tendency to jump to conclusions and to developed fixed beliefs more easily even in remission and when tasks are presented unrelated to delusional themes (e.g., Moritz and Woodward, 2005; Moritz et al., 2007). In the presence of psychotic-like experiences, persons seek less advice when making decisions (Scheunemann et al., 2021) and persons with a diagnosis of a psychotic disorder thought that a multitude of different interpretations of a given situation was plausible even when provided scarce or implausible explanations (Moritz and Woodward, 2004). Participants with psychotic disorder have also demonstrated both higher levels of certainty and a higher error rate in two studies on source attribution and the degree of subjective certainty for this judgement (Moritz and Woodward, 2002; Moritz et al., 2003). Furthermore, a bias against disconfirmatory evidence has been demonstrated repeatedly as an additional potential mechanism for the development and persistence of delusional ideation (Moritz and Woodward, 2006; Woodward et al., 2006a,b, 2008; Veckenstedt et al., 2011).

Beliefs seen through the lens of computational psychiatry

The new burgeoning field of "computational psychiatry" seeks to complement traditional, symptom-based diagnostic schemes in psychiatry with mathematical modeling in order to infer on the mechanisms which generate observed behavior and brain activity in psychiatric patients (Stephan and Mathys, 2014). Next to this theory- and mechanism-driven interpretation of computational psychiatry, another important trend has been to use so-called big data approaches and interrogate them in a pure data-driven manner by using machine learning algorithms (Huys et al., 2016). To address the topic of belief formation and updating in psychiatric conditions the theory-driven approach appears to be particularly well-suited, because it involves using mathematical models that formally describe the cognitive processes, including beliefs and their probabilities, that underlie observable behavior. To this end, a wide variety of models can be used, but two have found particularly widespread application: models of reinforcement learning and Bayesian inference. The latter approach has even give rise to the notion that the human brain can be described as a "Bayesian brain" that constructs and continuously updates a generative model of its sensory inputs (Knill and Pouget, 2004; Friston, 2010).

According to this approach, behavioral and neuroimaging studies are conducted that, for instance, use probabilistic learning tasks that ask study participants to learn from different types of information. In one such study conducted in our lab, we used a task that required learning about the winning probabilities of two cards and about the probability of a face giving the player advice by shifting gaze toward one of the two cards (Henco et al., 2020). Importantly, we did not explicitly tell participants to learn about the social information. The two types of information (non-social and social) were varied independently of each other during the course of the experiment, thereby constituting a volatile context, in which study participants with three major and severe psychiatric disorders were investigated: major depression (MDD; $n = 29$), schizophrenia (SCZ; $n = 31$) and borderline personality disorder (BPD; $n = 31$). In addition a group of participants was investigated without a history of a psychiatric disorder ($n = 34$). In other words, the study investigated whether volatility and probability learning is equally affected when inferring on the hidden states of non-social and social outcomes across the three different patient groups. We used the so-called hierarchical Gaussian filter (HGF; Mathys et al., 2011) to obtain a profile of each participant's particular way of updating beliefs when receiving social and non-social information while making decisions and selecting one of the cards on each trial. The HGF is a generic hierarchical Bayesian inference model for volatile environments with parameters that reflect individual variations in cognitive style. We went beyond other recent computational psychiatry studies using the HGF by using two parallel HGF hierarchies for social and non-social aspects of the environment. Our modeling framework was, thus, specifically designed to quantify the relative weight participants afforded their beliefs about the predictive value of social compared to non-social information. We found that patients with SCZ and BPD showed significantly poorer overall performance compared to healthy participants and patients with MDD, which raises the question which mechanisms underlie these patterns of behavior. Here, mathematical modeling allowed insights into how beliefs are updated and how these beliefs are translated into decisions: Results demonstrated revealed that SCZ and BPD patients both weighted their social-domain predictions more strongly than healthy study participants and patients MDD (Figure 1). This explains the lower performance of BPD and SCZ patients.

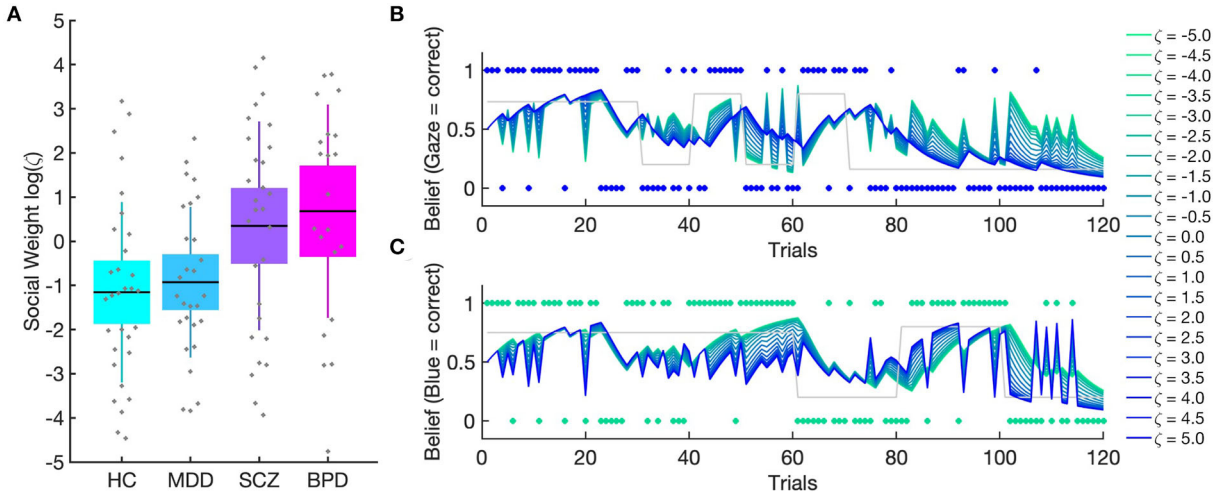


FIGURE 1
(Taken from: Henco et al., 2020). **(A)** Patients with borderline personality disorder (BPD) and schizophrenia (SCZ) gave social information more weight compared to healthy controls (HC) and patients with major depressive disorder (MDD). Boxes mark 95% confidence intervals and vertical lines standard deviations. **(B)** Simulations demonstrate the impact of varying weighting factor on combined belief. **(B)** shows that the combined belief of agents with high zeta values is aligned with the social input structure (blue dots) whereas these agents show a stochastic belief structure with regard to the non-social input structure (green dots) in **(C)**. Conversely, agents with low zeta values show a belief structure more closely aligned with the non-social input structure and a stochastic belief structure with regard to the social input as shown in **(B)**.

Their stronger reliance on social cues was detrimental because the social cue was more volatile than the non-social one. The commonality of over-weighting social-domain predictions in SCZ and BPD patients suggests itself as the decision-making aspect of a general interpersonal hypersensitivity in both conditions. This is also reflected in excessive, albeit inaccurate, mental state attributions (also described as hypermentalizing) that are often observed in patients with BPD and SCZ.

Hypermentalizing is also a possible explanation for the findings by Seow and Gillan (2020), who used similar modeling to show that healthy participants at the high end of the paranoia spectrum used similar weighting of social information irrespective of whether incorrect advice was framed to be intentional or not, while low-paranoia participants reduced their social weighting when negative advice was cued to be intentional. This indicates that dysfunctional belief systems also play a role in the normal population and may present a dimensional continuum, on which different individuals can be placed. Interestingly, environmental changes also seem to affect belief systems and belief updating: Research by Suthaharan et al. (2021) has demonstrated that the initial phase of the COVID pandemic in 2020 increased individuals' paranoia and made their belief updating more erratic. This was examined by combining self-rated paranoia scores and computerized social and non-social belief updating tasks. Here, it was found that the increase in self-rated paranoia was less pronounced in US states that enforced a more proactive lockdown and more pronounced at reopening in states that mandated mask-wearing. Computational modeling revealed that certain types of behavior (win-switch) and volatility priors tracked these changes in self-reported paranoia with policy.

Taken together, computational psychiatry provides important new tools to investigate the mechanisms that underlie cognition and behavior. By doing so, computational psychiatry aims at establishing a new and mathematically formalized way of assessing beliefs, how they change over time and how they relate to subjective experience and observable behavior. This powerful approach also holds great potential for the investigation of the neurophysiology related to psychiatric conditions and may inform differential diagnosis and subgroup detection in accordance with what has been described as personalized or precision psychiatry (Stephan and Mathys, 2014; Friston et al., 2017). In addition, it has more recently been suggested that computational psychiatry could also play an important role in elucidating the computational mechanisms of cognitive and behavioral psychotherapeutic interventions, which often aim at changing a person's beliefs in order to alleviate symptoms and treat psychiatric conditions (Moutoussis et al., 2018; Nair et al., 2020; Smith et al., 2021).

Discussion and outlook: How beliefs (and changing them) can help to treat psychiatric disorders

The cognitive turn of what at the time was still described as "behavioral therapy" consisted in introducing the idea that beliefs play an important role in mediating between so-called

"activating events" and "emotional consequences" (Ellis, 1958). In addition, the notion of "cognitive distortions" was introduced to indicate that beliefs can be unhelpful of even distorted (Beck, 1963). Such unhelpful beliefs can lead to negative emotions and maladaptive actions, thereby forming what has become known as a cognitive triangle (belief–affect–behavior). Reviewing the complete history of what is today referred to as "cognitive behavioral therapy" (CBT) and studies to investigate its effects and underlying mechanisms clearly is beyond the scope of this article. But it is safe to say that a large and increasing body of literature indicates that CBT techniques such as disputation of beliefs and cognitive restructuring are efficient and allow to target the general and specific belief systems that are deemed relevant for different psychiatric conditions. With regard to psychotic disorders, in particular, it has been demonstrated that meta-cognitive training (MCT) is a novel cognitive approach geared toward the treatment of positive symptoms in psychosis, but also other clinical conditions. MCT tries to help individuals experiencing psychosis to become more aware of the beliefs involved in their illness and to counteract the biased beliefs and assumptions that may predispose an individual to develop delusions (see Moritz et al., 2022 for a recent overview). Importantly, MCT typically takes place in a group setting, which allows for social interaction and exchange between different persons with the aim to reflect upon experiences and thoughts from different perspectives and consider information provided by others. In fact, the social interactions between group members are seen as a crucial aspect and key to the learning process.

Consistent with this, it is well-known that the social interaction between patient and therapist and the so-called therapeutic relationship plays a fundamental role in contributing to the success of psychotherapy (Leahy, 2008). In this regard, the concept of a so-called "complementary therapeutic relationship" has been proposed (Caspar et al., 2005), which suggests that therapists are "supposed to offer each patient an individually custom tailored relationship that suits his or her important goals." In other words, the therapist should adjust to the interactional profile and needs of the patient in order to contribute to and allow for a smooth and harmonious interaction, which lays the foundation for a helpful therapeutic relationship by promoting the development of trust. This also resonates with findings from non-clinical populations, where it has been demonstrated that the degree of interpersonal similarity is closely related to relationship quality (Bolis et al., 2021). In other words, how well/little people match interpersonally is important for the success of social interaction and communication, both in a non-clinical and a clinical context. This has been described as the "social interaction mismatch hypothesis" and can help to guide studies in social neuroscience toward the investigation of cross-brain processes (see Redcay and Schilbach, 2019 for a review). In a therapeutic context it is often a requirement that the therapist adjusts to the interactional needs of the patient to create a therapeutic relationship that later on can also be used to initiate change and to help the patient make so-called corrective experiences that challenge one's fears or expectations.

In addition, it can be argued that "disorders of social interaction" represent a defining feature of psychiatric disorders (Schilbach, 2016) and that addressing social interaction difficulties as a transdiagnostic phenomenon constitutes an important therapeutic goal (Schilbach et al., 2022).

With regard to the dyadic micro-processes relevant for the establishment of a helpful, motivating and trusting therapeutic relationship, it has been demonstrated that engaging in joint attention and sharing experiences with another person–even outside a therapeutic context – recruits reward-related neuro-circuitry, which can be interpreted in terms of an intrinsic motivation for social connection (Schilbach et al., 2010; Pfeiffer et al., 2014). Moreover, results from a wide range of studies demonstrate that non-verbal synchrony in dyadic interactions plays an important role to create rapport and may act as a "social glue" that binds persons together (Schilbach et al., 2008; Neufeld et al., 2016; for a review see Schilbach, 2015). Consistent with these ideas, results from a study by Ramseyer and Tschacher (2011) have shown that differences in synchrony at the level of nonverbal behavior are linked to relationship quality and therapy outcome (see also Koole and Tschacher, 2016 for a review). In this regard, the advent of novel technologies and methodologies now allows for a more fine-grained and quantitative analysis of interpersonal behavior during dyadic social interactions (Lahnakoski et al., 2020). Here, it has been found that it is not only synchrony that matters, but that aspects of interpersonal orienting and distance also predict the subjective quality of social interactions. Combining these new methods with computational modeling and other approaches from computational psychiatry promises to provide completely new insights into the mechanisms of social interaction and into how beliefs are (sometimes) shared across different brains (Henco and Schilbach, 2021). Furthermore, it appears feasible to investigate how differences in nonverbal synchrony during a social interaction may influence beliefs a person holds. Here, the dynamics of a social interaction could help to consolidate or change previously acquired social beliefs by providing a form of social feedback and validation. In combination with experimental tasks that record psychophysiology from two interacting persons, these developments could help to increase our understanding of how to improve the relationship quality and efficacy of psychotherapeutic interventions in the future (Bolis et al., 2017).

In summary, beliefs play a prominent role in our daily lives and help us to successfully navigate the environment and social interactions by providing probabilistic estimates of what we can hold to be true. But beliefs can also lead us astray and cause intense suffering as evident in psychotic disorders, but also a wide range of other psychiatric conditions. Fortunately, beliefs–in many, if not all instances–are subject to change or can be

recognized as just that, beliefs. Consequently, we may even show adaptive behavior in the presence of unhelpful beliefs and can make new experiences - often during social interactions - that may help us to leave certain beliefs behind.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Funding

This paper is funded by Rüdiger Seitz, via the Volkswagen Foundation, Siemens Healthineers, and the Betz Foundation.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Beck, A. T. (1963). *Thinking and depression:* idiosyncratic *content and cognitive distortions. Arch. Gen. Psychiatry.* 9, 324–333. doi: 10.1001/archpsyc.1963.01720160014002

Bolis, D., Balsters, J., Wenderoth, N., Becchio, C., and Schilbach, L. (2017). Beyond autism: introducing the dialectical misattunement hypothesis and a bayesian account of intersubjectivity. *Psychopathology* 50, 355–372. doi: 10.1159/000484353

Bolis, D., Lahnakoski, J. M., Seidel, D., Tamm, J., and Schilbach, L. (2021). Interpersonal similarity of autistic traits predicts friendship quality. *Soc. Cogn. Affect. Neurosci.* 16, 222–231. doi: 10.1093/scan/nsaa147

Caspar, F., Grossmann, C., Unmüssig, C., and Schramm, E. (2005). Complementary therapeutic relationship: therapist behavior, interpersonal patterns, and therapeutic effects. *Psychother. Res.* 15, 91–102. doi: 10.1080/10503300512331327074

Ellis, A. (1958). *Outcome of employing three techniques of psychotherapy. J. Clin. Psychol.* 13, 344–350.

Fletcher, P. C., and Frith, C. D. (2009). Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia. *Nat. Rev. Neurosci.* 10, 48–58. doi: 10.1038/nrn2536

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* 11, 127–138. doi: 10.1038/nrn2787

Friston, K. J., Redish, A. D., and Gordon, J. A. (2017). Computational nosology and precision psychiatry. *Comput Psychiatr.* 1, 2–23. doi: 10.1162/cpsy_a_00001

Hemsley, D. R., and Garety, P. A. (1986). The formation and maintenance of delusions: a Bayesian analysis. *Br. J. Psychiatry.* 149, 51–56. doi: 10.1192/bjp.149.1.51

Henco, L., Diaconescu, A. O., Lahnakoski, J. M., Brandi, M. L., Hörmann, S., Hennings, J., et al. (2020). Aberrant computational mechanisms of social learning and decision-making in schizophrenia and borderline personality disorder. *PLoS Comput. Biol.* 16, e1008162. doi: 10.1371/journal.pcbi.1008162

Henco, L., and Schilbach, L. (2021). Studying social inferences in and across social brains. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* 6, 760–761. doi: 10.1016/j.bpsc.2021.04.008

Huys, Q. J., Maia, T. V., and Frank, M. J. (2016). Computational psychiatry as a bridge from neuroscience to clinical applications. *Nat. Neurosci.* 19, 404–413. doi: 10.1038/nn.4238

Knill, D. C., and Pouget, A. (2004). The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci.* 27, 712–719. doi: 10.1016/j.tins.2004.10.007

Koole, S. L., and Tschacher, W. (2016). Synchrony in psychotherapy: a review and an integrative framework for the therapeutic alliance. *Front. Psychol.* 7, 862. doi: 10.3389/fpsyg.2016.00862

Lahnakoski, J. M., Forbes, P. A. G., McCall, C., and Schilbach, L. (2020). Unobtrusive tracking of interpersonal orienting and distance predicts the subjective quality of social interactions. *R. Soc. Open Sci.* 7, 191815. doi: 10.1098/rsos.191815

Leahy, R. L. (2008). The therapeutic relationship in cognitive-behavioral therapy. *Behav. Cogn. Psychother.* 36, 769–777. doi: 10.1017/S1352465808004852

Mathys, C., Daunizeau, J., Friston, K. J., and Stephan, K. E. (2011). A bayesian foundation for individual learning under uncertainty. *Front. Hum. Neurosci.* 5, 39. doi: 10.3389/fnhum.2011.00039

Moritz, S., Menon, M., Balzan, R., and Woodward, T. S. (2022). Metacognitive training for psychosis (MCT): past, present, and future. *Eur. Arch. Psychiatry Clin. Neurosci.* doi: 10.1007/s00406-022-01394-9

Moritz, S., and Woodward, T. S. (2002). Memory confidence and false memories in schizophrenia. *J. Nerv. Ment. Dis.* 190, 641–643. doi: 10.1097/00005053-200209000-00012

Moritz, S., and Woodward, T. S. (2004). Plausibility judgment in schizophrenic patients: evidence for a liberal acceptance bias. *Ger. J. Psychiatry.* 7, 66–74.

Moritz, S., and Woodward, T. S. (2005). Jumping to conclusions in delusional and non-delusional schizophrenic patients. *Br. J. Clin. Psychol.* 44, 193–207. doi: 10.1348/014466505X35678

Moritz, S., and Woodward, T. S. (2006). A generalized bias against disconfirmatory evidence in schizophrenia. *Psychiatry Res.* 142, 157–165. doi: 10.1016/j.psychres.2005.08.016

Moritz, S., Woodward, T. S., and Lambert, M. (2007). Under what circumstances do patients with schizophrenia jump to conclusions? A liberal acceptance account. *Br. J. Clin. Psychol.* 46, 127–137. doi: 10.1348/014466506X129862

Moritz, S., Woodward, T. S., and Ruff, C. C. (2003). Source monitoring and memory confidence in schizophrenia. *Psychol. Med.* 33, 131–139. doi: 10.1017/S0033291702006852

Moutoussis, M., Shahar, N., Hauser, T. U., and Dolan, R. J. (2018). Computation in psychotherapy, or how computational psychiatry can aid learning-based psychological therapies. *Comput Psychiatr.* 2, 50–73. doi: 10.1162/CPSY_a_00014

Nair, A., Rutledge, R. B., and Mason, L. (2020). Under the hood: using computational psychiatry to make psychological therapies more mechanism-focused. *Front. Psychiatry.* 11, 140. doi: 10.3389/fpsyt.2020.00140

Neufeld, J., Ioannou, C., Korb, S., Schilbach, L., and Chakrabarti, B. (2016). Spontaneous facial mimicry is modulated by joint attention and autistic traits. *Autism Res.* 9, 781–789. doi: 10.1002/aur.1573

Pfeiffer, U. J., Schilbach, L., Timmermans, B., Kuzmanovic, B., Georgescu, A. L., Bente, G., et al. (2014). Why we interact: on the functional role of the striatum in the subjective experience of social interaction. *Neuroimage* 101, 124–137. doi: 10.1016/j.neuroimage.2014.06.061

Ramseyer, F., and Tschacher, W. (2011). Nonverbal synchrony in psychotherapy: coordinated body movement reflects relationship quality and outcome. *J. Consult. Clin. Psychol.* 79, 284–295. doi: 10.1037/a0023419

Redcay, E., and Schilbach, L. (2019). Using second-person neuroscience to elucidate the mechanisms of social interaction. *Nat. Rev. Neurosci.* 20, 495–505. doi: 10.1038/s41583-019-0179-4

Scheunemann, J., Fischer, R., and Moritz, S. (2021). Probing the hypersalience hypothesis-an adapted judge-advisor system tested in individuals with psychotic-like experiences. *Front. Psychiatry*. 12, 612810. doi: 10.3389/fpsyt.2021.612810

Schilbach, L. (2015). Eye to eye, face to face and brain to brain: novel approaches to study the behavioral dynamics and neural mechanisms of social interactions. *Curr. Opin. Behav. Sci.* 3, 130–135. doi: 10.1016/j.cobeha.2015.03.006

Schilbach, L. (2016). Towards a second-person neuropsychiatry. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 371, 20150081. doi: 10.1098/rstb.2015.0081

Schilbach, L., Bartholomäus, M., Böhm, J., Buchholz, F., Parpart, H., Albantakis, L., et al. (2022). *Kompetenz-orientiertes Manual zur Behandlung von Störungen der sozialen Interaktion*. Elsevier.

Schilbach, L., Eickhoff, S. B., Mojzisch, A., and Vogeley, K. (2008). What's in a smile? Neural correlates of facial embodiment during social interaction. *Soc Neurosci.* 3, 37–50. doi: 10.1080/17470910701563228

Schilbach, L., Wilms, M., Eickhoff, S. B., Romanzetti, S., Tepest, R., Bente, G., et al. (2010). Minds made for sharing: initiating joint attention recruits reward-related neurocircuitry. *J. Cogn. Neurosci.* 22, 2702–2715. doi: 10.1162/jocn.2009.21401

Seitz, R. (2022). Beliefs: a challenge in neuropsychological disorders. *J. Neuropsychol.* 16, 21–37. doi: 10.1111/jnp.12249

Seow, T. X. F., and Gillan, C. M. (2020). Transdiagnostic phenotyping reveals a host of metacognitive deficits implicated in compulsivity. *Sci. Rep.* 10, 2883. doi: 10.1038/s41598-020-59646-4

Smith, R., Moutoussis, M., and Bilek, E. (2021). Simulating the computational mechanisms of cognitive and behavioral psychotherapeutic interventions: insights from active inference. *Sci. Rep.* 11, 10128. doi: 10.1038/s41598-021-89047-0

Stephan, K. E., and Mathys, C. (2014). Computational approaches to psychiatry. *Curr. Opin. Neurobiol.* 25, 85–92 doi: 10.1016/j.conb.2013.12.007

Suthaharan, P., Reed, E. J., Leptourgos, P., Kenney, J. G., Uddenberg, S., Mathys, C. D., et al. (2021). Paranoia and belief updating during the COVID-19 crisis. *Nat Hum Behav.* 5, 1190–1202. doi: 10.1038/s41562-021-01176-8

Veckenstedt, R., Randjbar, S., Vitzthum, F., Hottenrott, B., Woodward, T. S., and Moritz, S. (2011). Incorrigibility, jumping to conclusions, and decision threshold in schizophrenia. *Cogn. Neuropsychiatry*. 16, 174–192. doi: 10.1080/13546805.2010.536084

Woodward, T. S., Moritz, S., and Chen, E. Y. (2006a). The contribution of a cognitive bias against disconfirmatory evidence (BADE) to delusions: a study in an Asian sample with first episode schizophrenia spectrum disorders. *Schizophr. Res.* 83, 297–298. doi: 10.1016/j.schres.2006.01.015

Woodward, T. S., Moritz, S., Cuttler, C., and Whitman, J. C. (2006b). The contribution of a cognitive bias against disconfirmatory evidence (BADE) to delusions in schizophrenia. *J. Clin. Exp. Neuropsychol.* 28, 605–617. doi: 10.1080/13803390590949511

Woodward, T. S., Moritz, S., Menon, M., and Klinge, R. (2008). Belief inflexibility in schizophrenia. *Cogn. Neuropsychiatry*. 13, 267–277. doi: 10.1080/13546800802099033

frontiers | Frontiers in Behavioral Neuroscience

TYPE Opinion
PUBLISHED 20 October 2022
DOI 10.3389/fnbeh.2022.956307

Check for updates

OPEN ACCESS

EDITED BY
Rüdiger J. Seitz,
Heinrich Heine University of Düsseldorf, Germany

REVIEWED BY
Motoaki Sugiura,
Tohoku University, Japan

*CORRESPONDENCE
Katharina Voigt
katharina.voigt@monash.edu

SPECIALTY SECTION
This article was submitted to Emotion Regulation and Processing, a section of the journal Frontiers in Behavioral Neuroscience

RECEIVED 30 May 2022
ACCEPTED 11 August 2022
PUBLISHED 20 October 2022

CITATION
Voigt K (2022) Where do our preferences come from? How hard decisions shape our preferences. *Front. Behav. Neurosci.* 16:956307. doi: 10.3389/fnbeh.2022.956307

COPYRIGHT
© 2022 Voigt. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Where do our preferences come from? How hard decisions shape our preferences

Katharina Voigt[1,2]*

[1]Turner Institute for Brain and Mental Health, Monash University, Clayton, VIC, Australia, [2]Monash Biomedical Imaging, Monash University, Clayton, VIC, Australia

KEYWORDS

preference, decision-making, choice-induced preference change, preference formation, value-based decision model

Introduction

Where do our preferences come from? Traditional neurocognitive models of value-based choice view decision-making as a serial process in which stable preferences are the basis of subsequent choices (Dolan and Dayan, 2013). An alteration of preferences is only expected if new (external) information about choice alternatives becomes available (e.g., through the consumption of a good). Accordingly, in a supermarket we assign values to items based on our stable preferences and choose the item we assigned the highest value to. After we tasted our selection, we can adjust our preferences for that item based on this recent experience. However, one highly debated question over the past decade has been whether preferences can change endogenously, that is, in the absence of any additional external information about the choice options, and merely as a function of our past choice history. Specifically, a growing body of studies found that when individuals must make binary choices between items they initially indicated to prefer equally well, their preferences for the chosen option increases and decreases for the rejected option. This empirical observation is now commonly referred to as the choice-induced preference change effect (reviewed by Izuma and Murayama, 2013; Enisman et al., 2021).

Prominent explanations of the choice-induced preference change effect are based on (Festinger, 1957) theory of cognitive dissonance, which proposes that discrepancies between actions and preferences cause psychological discomfort. Preferences are then adjusted *after* a hard decision has been made to reduce the dissonance between initial preference and the decision outcome (reviewed by Harmon-Jones et al., 2015). This explanation is in line with neuroimaging studies, which suggested that at the time of re-evaluation, after dissonance between preferences and choices is detected by the anterior cingulate cortex (ACC; van et al., 2009; Kitayama et al., 2013), the dorsolateral prefrontal cortex (dlPFC) triggers changes in the neural representation of value (Izuma et al., 2010, 2015; Mengarelli et al., 2015) in the ventromedial PFC (vmPFC) or ventral striatum (vStr; Izuma et al., 2010; Chammat et al., 2017). However, what happens in situations when we equally prefer two choice alternatives and therefore existing preferences are not sufficient to differentiate among them? In other words, how do we solve hard decisions?

An alternative possibility is that preferences are adjusted much earlier, that is, while a hard decision is being made, when the value differential of the options is not sufficient to choose among them. As such, preference adjustments might constitute a necessary adaptive (online) mechanism to deal with hard choices, as opposed to a post-decisional process for eliminating cognitive dissonance (Izuma et al., 2010, 2015). This new hypothesis, however, remains largely untested as existing functional neuroimaging studies focused entirely on the neural mechanisms of preference change during re-evaluation (Izuma et al., 2010; Chammat et al., 2017). Studying decisions among equally preferred items, however, holds key in understanding how our preferences are dynamically constructed based on the choice context.

Our recent neuroimaging study is the first to study preferences changes during hard decisions and provides evidence for this alternative theory of choice-induced preference changes (Voigt et al., 2019). Preference changes were predicted from activity in left dorsolateral prefrontal cortex and precuneus while making hard decisions. Fixation durations during this phase predicted both choice outcomes and subsequent preference changes. These preference adjustments became behaviorally relevant only for choices that were remembered and were in turn associated with hippocampus activity. These findings suggest that preferences evolve dynamically as decisions arise, potentially as a mechanism to prevent stalemate situations in underdetermined decision scenarios. Based on this recent evidence from neuroimaging I propose a novel neural framework of choice-induced preference changes, which is described in the following.

An integrative neural model underlying endogenous preference formation during hard decisions

In the next section i will integrate recent empirical evidence that supports an alternative model of choice-induced preference effects. This tentative, integrative neural process model describes endogenous preference formation during hard decisions in four processing steps: (1) value computation (2) conflict detection, (2) value updating, (3) value-based decision, (4) updated value representation (Figure 1).

(1) Value computation: The first process involves the computation of value for both choice alternatives. This computation facilitates a comparative process, allowing the decision maker to identify and pursue the option with greatest expected value (Samuelson, xbib1938). Value computation is further essential in triggering the preparation of upcoming motor responses: to make appropriate decisions, these values must be reliable predictors of the benefits that are likely to result from each action. There is converging evidence that value computation is predominantly associated with prefrontal areas, such as the vmPFC (Kable and Glimcher, 2007; Chib et al., 2009; Bartra et al., 2013), dlPFC (Hare et al., 2009; Sokol-Hessner et al., 2013), but also the vStr (Bartra et al., 2013). The vmPFC in particular has been shown to compute value as a 'common currency' (Chib et al., 2009) and its activity is related to upcoming computations of motor choice responses (Rudorf and Hare, 2014).

(2) Conflict detection: If the result of value computation is that the value differential between both items is not sufficient to distinguish between the choice items, and therefore, no motor preparations can be triggered, a moment of indecision or decision conflict occurs. If no decision conflict is detected, that is, the value differential is sufficient to discriminate among the options (e.g., during easy decisions), then the system can compute a decision value and continues with the value-based choice (as described in the fourth step of this model). Behavioral evidence suggests that preference changes do not occur for easy choices, but only for hard decisions and individuals take significantly longer to solve hard over easy decisions (Voigt et al., 2019). Neuroimaging studies revealed that hard decisions, compared to easy decisions, were associated with dACC activity (Kitayama et al., 2013; Voigt et al., 2019). Previous studies linked activity in the dACC to decision conflicts (Kitayama et al., 2013; Shenhav et al., 2013; Shenhav and Buckner, 2014).

3) Value updating: Fixation duration plays a causal role in value-guided choice (e.g., Shimojo et al., 2003; Glaholt et al., 2009). Krajbich et al. (2010) showed that the fixation duration for an item mirrors the evidence accumulation process for an decision outcome. Eye-tracking data during the process of making hard decisions revealed that the total and first fixation duration for of an item was predictive of its choice and, importantly, its subsequent change in value (when controlling for choice) (Voigt et al., 2019). The latter finding gives reason to assume that fixations contribute to the construction process of new preference values prior to the choice. In the light of these and our findings, the proposed preference formation model implies that in underdetermined decision scenarios, the decision system extracts new information in the moment of choice via fixation on a particular item in order to construct new preferences guiding upcoming choices. This in turn might reflect an adaptive mechanism to solve hard decisions.

Activity in the dlPFC and precuneus was linked with online, trial-by-trial updates of preferences during hard choices (Voigt et al., 2019). The left dlPFC was previously shown to be involved in the implementation of preference change *after* the difficult choice was made (Izuma et al., 2010; Harmon-Jones et al., 2015; Mengarelli et al., 2015). Both the dlPFC and precuneus have been previously associated with dissociable roles in working memory (Brodt et al., 2016), shifts of spatial attention (Yan et al., 2016) and value reconstruction (Harris et al., 2011). Specifically, the precuneus was shown to be involved in the early bottom-up selection of spatial attention, whereby the dlPFC was associated with later top-down selection of spatial attention. As

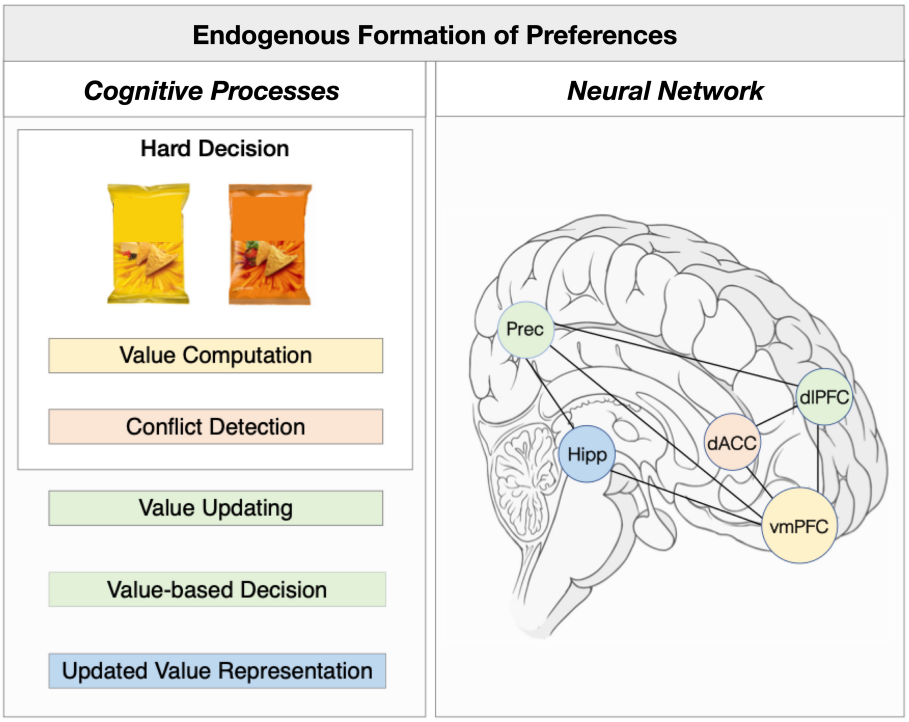


FIGURE 1
Integrative neural model underlying endogenous preference formation. Left Panel: The underlying cognitive mechanisms of choice-induced preference change effects can be described into four main processing steps: (1) value computation (orange) (2) conflict detection (red), (2) value updating and (3) value-based decision (green), (4) updated value representation (blue). Right Panel: Neural regions and networks associated with the cognitive processing steps of choice-induced preference change effects. Colors of cognitive processes match the neural correlates as depicted in the neural processes. Hipp, Hippocampus; Prec, Precuneus; dACC, dorsal anterior cingulate cortex; vmPFC, ventromedial prefrontal cortex; dlPFC, dorsolateral prefrontal cortex.

fixations were shown to play a role in preference reconstruction it is reasonable to assume that initially the precuneus, which has rich connections to the superior colliculus administering eye-movements (Yeterian and Pandya, 1998), allocates spatial attention to the salient stimulus (i.e., bottom up) and forwards this information to the dlPFC. Previous studies showed that the dlPFC 'holds' choice-relevant information in working memory (Brodt et al., 2016) in order to guide performances toward targets. This temporal dynamic value representation evolving from posterior to prefrontal regions was demonstrated in other studies (Harris et al., 2011). These findings suggest that initial bottom-up shifts in spatial attention are explained by in-decisional precuneus activity controlling the reconstruction of new value information, which is forwarded to the dlPFC. This representation is then stored into working memory assisting the individual to maximally discriminate between the choice options and implementing the choice.

(4) Value-based decision: Based on the rapid fixation-guided computation of a new value signal, which might be subserved by a network subtending the precuneus and dlPFC, the decision system is now able to compute a decision variable as the value differential among the alternatives is now sufficient to distinguish between them. This means, the decision scenario now has transferred from a hard to a (relatively) easy one. The computation of decision value has been associated with dlPFC activity, which is connected to premotor areas conducting the actual motor choice (Miller and Cohen, 2001). Further, transient disruption of human dlPFC induced by theta-burst transcranial magnetic stimulation has been shown to interfere with forward planning and flexible, outcome-specific decision behavior (Smittenaar et al., 2013). This functionality of dlPFC activity might also explain why previous research investigating choice-induced preference change effects found the involvement of the dlPFC at the stage of re-evaluation (e.g., Izuma et al., 2010, 2015). In the model, this might simply reflect the computation of an upcoming value-based choice, but not necessarily the update of value itself as it was previously proposed.

(5) Updated value representation: Although preference changes were updated during hard choices, behavioral evidence indicates that only choice outcomes that are explicitly remembered were encoded as preference changes long-term (i.e., at re-evaluation); choice outcomes that were forgotten or guessed did not trigger long-lasting preference changes (Salti et al., 2014; Chammat et al., 2017; Voigt et al., 2019).

Neuroimaging studies showed that this memory-dependent choice-induced preference change effect is associated with left hippocampus activity (Chammat et al., 2017; Voigt et al., 2019). The hippocampus plays a key role in long-term and it has strong reciprocal anatomical connections with the vmPFC (Weilbächer and Gluth, 2017), supporting its role in long-term representations of endogenous preference changes.

Conclusion

This opinion paper presented empirical findings that support the notion that our preferences evolve endogenously during the process of making decisions between equally preferred items. In those situations of indecision, the information gathered via fixation toward an item seems critical to reconstruct upcoming value-based decisions. This online mechanism of fixation-driven preference formation might be depended on the idiosyncrasy of an underlying prefrontal-parietal brain network. Such rapid changes in preferences prior to the initial undetermined decision, could constitute an adaptive mechanism enabling the individual to act. These in-decisional changes in preferences become behaviorally manifested when choice outcomes were explicitly remembered and encoded by episodic memory regions. Overall, these findings suggest a potential rethinking of the very notion of preferences and value-based choice. Further, they suggest a shift away from previous explanations of endogenous preference formation. Rather these findings suggest that seemingly self-determined, subjective cognitive concepts, such as our preferences, might be emergent consequences from the particulars of the decision scenario itself and brain networks underlying value-based decisions.

Author contributions

The author confirms being the sole contributor of this work and has approved it for publication.

Funding

This paper was funded by Dr. Rüdiger Seitz, via the Volkswagen Foundation, Siemens Healthineers, and the Betz Foundation.

Acknowledgments

The author thanks A/Prof. Stefan Bode and Prof. Carsten Murawski for their valuable and essential support for the presented research work.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Bartra, O., McGuire, J. T., and Kable, J. W. (2013). The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *NeuroImage* 76, 412–427. doi: 10.1016/j.neuroimage.2013.02.063

Brodt, S., Pöhlchen, D., Flanagin, V. L., Glasauer, S., Gais, S., and Schönauer, M. (2016). Rapid and independent memory formation in the parietal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 113, 13251–13256. doi: 10.1073/pnas.1605719113

Chammat, M., Karoui, I. E., Allali, S., Hagège, J., Lehongre, K., Hasboun, D., et al. (2017). Cognitive dissonance resolution depends on episodic memory. *Sci. Rep.* 7:41320. doi: 10.1038/srep41320

Chib, V. S., Rangel, A., Shimojo, S., and O'Doherty, J. P. (2009). Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *J. Neurosci.* 29, 12315. doi: 10.1523/JNEUROSCI.2575-09.2009

Dolan, R. J., and Dayan, P. (2013). Goals and habits in the brain. *Neuron* 80, 312–325. doi: 10.1016/j.neuron.2013.09.007

Enisman, M., Shpitzer, H., and Kleiman, T. (2021). Choice changes preferences, not merely reflects them: a meta-analysis of the artifact-free free-choice paradigm. *J. Personal. Soc. Psychol.* 120, 16–29. doi: 10.1037/pspa0000263

Festinger, L. (1957). *A Theory of Cognitive Dissonance*. Redwood City, CA: Stanford University Press.

Glaholt, M. G., Wu, M.-C., and Reingold, E. M. (2009). Predicting preference from fixations. *PsychNology J.* 7, 141–158.

Hare, T. A., Camerer, C. F., and Rangel, A. (2009). Self-control in decision-making involves modulation of the vmPFC valuation system. *Science* 324, 646–648. doi: 10.1126/science.1168450

Harmon-Jones, E., Harmon-Jones, C., and Levy, N. (2015). An action-based model of cognitive-dissonance processes. *Curr. Direct. Psychol. Sci.* 24, 184–189. doi: 10.1177/0963721414566449

Harris, A., Adolphs, R., Camerer, C., and Rangel, A. (2011). Dynamic construction of stimulus values in the ventromedial prefrontal cortex. *PLoS ONE* 6:13. doi: 10.1371/journal.pone.0021074

Izuma, K., Akula, S., Murayama, K., Wu, D.-A., Iacoboni, M., and Adolphs, R. (2015). A causal role for posterior medial frontal cortex

in choice-induced preference change. *J. Neurosci.* 35, 3598–3606. doi: 10.1523/JNEUROSCI.4591-14.2015

Izuma, K., Matsumoto, M., Murayama, K., Samejima, K., Sadato, N., and Matsumoto, K. (2010). Neural correlates of cognitive dissonance and choice-induced preference change. *Proc. Natl. Acad. Sci. U.S.A.* 107, 22014–22019. doi: 10.1073/pnas.1011879108

Izuma, K., and Murayama, K. (2013). Choice-induced preference change in the free-choice paradigm: a critical methodological review. *Front. Psychol.* 4, 41. doi: 10.3389/fpsyg.2013.00041

Kable, J. W., and Glimcher, P. W. (2007). The neural correlates of subjective value during intertemporal choice. *Nat. Neurosci.* 10, 1625–1633. doi: 10.1038/nn2007

Kitayama, S., Chua, H. F., Tompson, S., and Han, S. (2013). Neural mechanisms of dissonance: An fMRI investigation of choice justification. *NeuroImage* 69, 206–212. doi: 10.1016/j.neuroimage.2012.11.034

Krajbich, I., Armel, C., and Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nat. Neurosci.* 13, 1292–1298. doi: 10.1038/nn.2635

Mengarelli, F., Spoglianti, S., Avenanti, A., and di Pellegrino, G. (2015). Cathodal tDCS over the left prefrontal cortex diminishes choice-induced preference change. *Cerebral Cortex* 25, 1219–1227. doi: 10.1093/cercor/bht314

Miller, E. K., and Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Ann. Rev. Neurosci.* 24, 167–202. doi: 10.1146/annurev.neuro.24.1.167

Rudorf, S., and Hare, T. A. (2014). Interactions between dorsolateral and ventromedial prefrontal cortex underlie context-dependent stimulus valuation in goal-directed choice. *J. Neurosci.* 34, 15988–15996. doi: 10.1523/JNEUROSCI.3192-14.2014

Salti, M., El Karoui, I., Maillet, M., and Naccache, L. (2014). Cognitive dissonance resolution is related to episodic memory. *PLoS ONE* 9:e108579. doi: 10.1371/journal.pone.0108579

Samuelson, P. A. (1938). A note on pure theory of consumer's behaviour. *Economica.* 5, 61–71.

Shenhav, A., Botvinick, M. M., and Cohen, J. D. (2013). The expected value of control: an integrative theory of anterior cingulate cortex function. *Neuron* 79, 217–240. doi: 10.1016/j.neuron.2013.07.007

Shenhav, A., and Buckner, R. L. (2014). Neural correlates of dueling affective reactions to win–win choices. *Proc. Natl. Acad. Sci. U.S.A.* 111, 10978–10983. doi: 10.1073/pnas.1405725111

Shimojo, S., Simion, C., Shimojo, E., and Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nature Neuroscience* 6:1317–1322. doi: 10.1038/nn1150

Smittenaar, P., FitzGerald, T. H. B., Romei, V., Wright, N. D., and Dolan, R. J. (2013). Disruption of dorsolateral prefrontal cortex decreases model-based in favor of model-free control in humans. *Neuron* 80, 914–919. doi: 10.1016/j.neuron.2013.08.009

Sokol-Hessner, P., Camerer, C. F., and Phelps, E. A. (2013). Emotion regulation reduces loss aversion and decreases amygdala responses to losses. *Soc. Cogn. Affect. Neurosci.* 8:341. doi: 10.1093/scan/nss002

van, V., Mk, K., Jw, S., and Cs, C. (2009). Neural activity predicts attitude change in cognitive dissonance. *Nat. Neurosci.* 12, 1469–1474. doi: 10.1038/nn.2413

Voigt, K., Murawski, C., Speer, S., and Bode, S. (2019). Hard decisions shape the neural coding of preferences. *J. Neurosci.* 39, 718–726. doi: 10.1523/JNEUROSCI.1681-18.2018

Weilbächer, R. A., and Gluth, S. (2017). The interplay of hippocampus and ventromedial prefrontal cortex in memory-based decision making. *Brain Sciences* 7:4. doi: 10.3390/brainsci7010004

Yan, Y., Wei, R., Zhang, Q., Jin, Z., and Li, L. (2016). Differential roles of the dorsal prefrontal and posterior parietal cortices in visual search: a TMS study. *Scientific Reports* 6:30300. doi: 10.1038/srep30300

Yeterian, E. H., and Pandya, D. N. (1998). Corticostriatal connections of the superior temporal region in rhesus monkeys. *J. Comparat. Neurol.* 399, 384–402. doi: 10.1002/(SICI)1096-9861(19980928)399:3andlt;384::AID-CNE7andgt;3.0.CO;2-X

frontiers | Frontiers in Behavioral Neuroscience

OPINION
published: 01 July 2022
doi: 10.3389/fnbeh.2022.943419

Neural Mechanisms Underlying Expectation-Guided Decision-Making

*Thorsten Kahnt**

National Institute on Drug Abuse Intramural Research Program, National Institutes of Health, Baltimore, MD, United States

Keywords: decision-making, expectation, outcome-guided, inference, orbitofrontal cortex, model-based

INTRODUCTION

Many decisions are guided by expectations about their outcomes. For instance, we may decide to visit a restaurant because we anticipate the food to be outstanding. How these expectations are represented in the brain, and how they allow us to make adaptive choices are important questions for understanding the neural basis of behavior.

Work across species has revealed brain areas that signal expected rewards (Haber and Knutson, 2010; Kahnt, 2018). This work typically focuses on neural correlates of the value of choice options (Padoa-Schioppa, 2011), that is, how desirable an option is. Activity in many brain areas, including the striatum, ventromedial prefrontal cortex and orbitofrontal cortex (OFC), is correlated with expected value. However, expected outcomes are more than their value—they have a specific identity. Even though we may equally desire pizza funghi and spaghetti arrabiata, they are not the same, and representing expectations about the identity of outcomes is important for adaptive decision-making.

In this opinion, I will summarize recent work from my lab that has shown how the lateral OFC represents expectations about specific outcomes, how these expectations are learned, and how they can be used for adaptive decision-making. Finally, I will summarize evidence that disrupting activity in OFC networks that represent specific outcome expectations impairs adaptive behavior. Together, these findings support the view that the OFC contributes to expectation-guided decision-making by enabling us to simulate the consequences of our choices.

NEURAL REPRESENTATIONS OF OUTCOME EXPECTATIONS

Recent studies have shown that the OFC represents not only expectations about the value of future outcomes but also their identity (Howard and Kahnt, 2021). For instance, in one study, we used food odors as specific rewards and selected one sweet and one savory odor for each subject that were matched in rated pleasantness (i.e., value) (Howard et al., 2015). We then lowered the concentration of the food odors to create a set of low-intensity odors, which were rated as less pleasant than the high-intensity odors. The four food odors were then paired with different visual stimuli, such that each odor was reliably predicted by a different symbol. Finally, subjects were presented with these symbols while undergoing functional magnetic resonance imaging (fMRI). Multi-voxel pattern analysis (Kahnt, 2018) to the fMRI responses evoked by the symbols revealed that activity patterns in the lateral OFC, anterior cingulate cortex, and hippocampus differentiated between the two expected food odors, whereas activity patterns in the medial OFC represented the value of the odors, independent of their identity. These findings are in line with other work from our lab (Howard and Kahnt, 2017) as well as with studies showing that activity patterns in the lateral OFC represent values that are tied to specific reward categories, whereas activity in the medial OFC is independent of reward category (Mcnamee et al., 2013).

OPEN ACCESS

Edited by:
Rüdiger J. Seitz, Heinrich Heine University of Düsseldorf, Germany

Reviewed by:
Serge H. Ahmed, Centre National de la Recherche Scientifique (CNRS), France

***Correspondence:**
Thorsten Kahnt
thorsten.kahnt@nih.gov

Specialty section:
This article was submitted to Emotion Regulation and Processing, a section of the journal Frontiers in Behavioral Neuroscience

Received: *13 May 2022*
Accepted: *14 June 2022*
Published: *01 July 2022*

Citation:
Kahnt T (2022) Neural Mechanisms Underlying Expectation-Guided Decision-Making. Front. Behav. Neurosci. 16:943419. doi: 10.3389/fnbeh.2022.943419

LEARNING OF OUTCOME EXPECTATIONS

Outcome expectations are based on associations between predictive stimuli and rewards, and these associations need to be learned and updated through experience. Work in non-human primates has shown that dopamine neurons in the midbrain contribute to learning the value of rewards by signaling reward prediction errors, or the difference between received and expected rewards (Schultz et al., 1997). We hypothesized that midbrain activity encodes a similar signal for identity prediction errors, which may be used for learning reward identity expectations.

In one experiment, hungry subjects were presented with visual symbols that predicted one of two preference-matched food odors (e.g., strawberry or potato chips) in either low or high intensity (Howard and Kahnt, 2018). As in previous studies, subjects reported a higher preference for the high-intensity odors, but there was no preference difference between the sweet and savory food odors. After a number of trials of receiving the predicted odor, either the identity (e.g., subjects expected strawberry but received equally-preferred potato chips) or the intensity (e.g., subjects expected potato chips in low intensity but received the preferred high-intensity odor) of the odor was unexpectedly changed. fMRI activity in the midbrain showed signatures of value-based prediction errors, increasing when subjects received the more preferred high-intensity odor after expecting the less preferred low-intensity odor. However, activity in the same midbrain region also increased when subjects received strawberry after expecting potato chips, in line with the signaling of value-neutral identity prediction errors. Importantly, value- and identity-based prediction errors were found in the same part of the midbrain and were correlated, suggesting that they may originate from the same neural population. Similar findings have been observed in a study that recorded activity from dopamine neurons in rats (Takahashi et al., 2017), as well as in other human imaging studies (Boorman et al., 2016; Schwartenbeck et al., 2016; Suarez et al., 2019).

A question that follows is whether midbrain identity prediction errors actively shape identity learning in downstream areas, or whether they merely act as a permissive gating (i.e., salience) signal to direct attention and boost learning (Bromberg-Martin et al., 2010). We addressed this question, reasoning that if identity prediction errors conveyed salience information without providing specific information, there should be no difference between the midbrain response to reward B when A was predicted and the midbrain response to reward A when B was predicted. In contrast, if identity prediction errors actively shape learning in downstream targets, they should contain specific information such that midbrain responses differ between these two cases. In line with the latter idea, we found that midbrain fMRI patterns in humans and dopamine ensemble responses in rats contain information about the specific identity of the error (Stalnaker et al., 2019), suggesting they could directly update identity expectations in downstream areas, such as OFC.

Indeed, we found that the magnitude of identity prediction error response in the midbrain was correlated with how much identity expectations in the lateral OFC changed after an identity error (Howard and Kahnt, 2018). This suggests that identity expectations in the lateral OFC are updated through a mechanism that involves identity prediction errors in the dopaminergic midbrain.

USING EXPECTATIONS FOR INFERENCE

In many cases, we can learn the expected value of choice options through direct experience. For instance, we can learn the value of an item on a restaurant menu by ordering it. However, for many other decisions in life, we simply have not had the opportunity to directly learn values in this way. This especially applies to decisions that are less frequently or only indirectly experienced, like deciding to try out a new restaurant or whether to visit a new country. Also, the values we have learned from previous decisions may have changed since we last made that choice, and using these old values would lead to maladaptive decisions. In these situations, value expectations need to be computed by mentally simulating or inferring the value of the option based on incomplete information. Specific outcome expectations allow us to do this because they are part of a model of the relevant environment which we can use to simulate the consequences of our actions.

Such simulations can be studied in the devaluation task. In a typical experiment, subjects first learn to associate different sensory cues with different foods, e.g., M&Ms and peanuts (Rudebeck et al., 2013; Murray et al., 2015; Reber et al., 2017). After one of the rewards is devalued by feeding the food to satiety, subjects can make choices between the sensory cues. To access the current value of the choice option, subjects must simulate what outcome they will receive by making a particular choice and infer its current value. This allows them to avoid selecting the cue that predicts the devalued outcome. In contrast, if they use the previously learned value, they will make choices that result in both the valued and the devalued outcome.

We have used transcranial magnetic stimulation (TMS) to test whether outcome identity expectations represented in the lateral OFC are necessary for adaptive responding in the devaluation task (Howard et al., 2020). Hungry participants first learned associations between visual symbols and sweet or savory food odors and were then allowed to make choices between these symbols. Stimulation coordinates in the lateral PFC were selected for each participant based on resting-state fMRI connectivity with lateral OFC. After a session of continuous theta burst stimulation (cTBS), which has inhibitory after-effects lasting for 50–60 min (Huang et al., 2005), or sham stimulation, subjects ate a meal that was matched to either the sweet or the savory food odor. After this devaluation procedure, subjects could again make choices between the cues. Targeting the lateral OFC with cTBS had profound effects on subjects' choices after the meal. Whereas, subjects in the sham group adaptively stopped selecting symbols that predicted the devalued odor, subjects in the cTBS group continued to select these stimuli. This shows that OFC activity is required for using specific outcome expectations for making inferences about the current value of choice options.

A different type of inference can be probed in the sensory preconditioning task (Brogden, 1939; Hoffeld et al., 1960). In this task, subjects first learn associations between sensory stimuli A and B, and C and D (A → B, C → D). Then, the second cue of each pair (B and D) is paired with either a reward or no reward (B → reward, D → no reward). Finally, responses to all stimuli (A, B, C, and D) are probed. Humans and other animals show stronger responding to stimulus A compared to stimulus C in this final test (Sadacca et al., 2016; Sharpe et al., 2017; Wang et al., 2020b). This pattern of responding is compatible with the idea that subjects mentally step through the associations A → B and B → reward to infer that A → reward.

Activity in the OFC correlates with learning of the stimulus-stimulus associations during the initial learning phase (Sadacca et al., 2018; Wang et al., 2020b), suggesting that the OFC represents the associative structure of the task. In other words, stimulus-stimulus associations appear to be represented in the same way as associations between a sensory stimulus and a food reward. Moreover, OFC is critical for using these associations to perform mental simulations. Pharmacological inactivation of the lateral OFC in rats (Jones et al., 2012) as well as cTBS targeting the lateral OFC network in humans before the final phase of the sensory preconditioning task impairs responding to cue A, without affecting responding to cue B (for which subjects had directly learned the stimulus-outcome associations) (Wang et al., 2020a). Thus, just like neural representations of specific outcome expectations, representations of stimulus-stimulus associations in the lateral OFC network are critical for making mental simulations required for adaptive decision-making.

DISCUSSION

The work described above outlines the neural mechanisms underlying expectation-guided decision-making. In brief, the OFC represents expectations about specific outcomes, and these expectations are learned through an error-based mechanism that involves the dopaminergic midbrain. The same networks that represent outcome expectations also represent expectations about future events, even if they do not possess any value. Of note, while we often make decisions between options with outcomes that belong to very different categories, our experiments used outcomes from the same reward category (i.e., food). This can be considered a stronger test of the outcome-specific coding hypothesis, because differences in neural responses to different reward categories may not only reflect outcome-specific coding but also different preparatory or consummatory reward responses. Thus, results from within category experiments are likely to generalize to across category settings. Indeed, previous work on neural representations of different reward categories has revealed comparable findings (Levy and Glimcher, 2011; Mcnamee et al., 2013; Gross et al., 2014).

Neural representations of specific outcomes enable us to perform mental simulations that are required for adaptive behavior in novel situations or when the value of an outcome has changed since we last made that decision. In other words, these representations allow us to flexibly assign value or meaning to expected outcomes in order to guide our decisions. Together, the findings discussed here are compatible with the view that the OFC network contributes to decision-making by representing a model of the environment, which enables us to make flexible inferences about the outcomes of our decisions.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

This study was funded by Dr. Rüdiger Seitz, *via* the Volkswagen Foundation, Siemens Healthineers, and the Betz Foundation. The author was supported by the Intramural Research Program at the National Institute on Drug Abuse, and the expressed opinions are the author's own and do not reflect the view of the NIH/DHHS. Siemens Healthineers was not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

ACKNOWLEDGMENTS

The author thanks the members of the Kahnt Lab who have performed the work summarized here.

REFERENCES

Boorman, E. D., Rajendran, V. G., O'reilly, J. X., and Behrens, T. E. (2016). Two anatomically and computationally distinct learning signals predict changes to stimulus-outcome associations in hippocampus. *Neuron* 89, 1343–1354. doi: 10.1016/j.neuron.2016.02.014

Brogden, W. J. (1939). Sensory pre-conditioning. *J. Exp. Psychol.* 25, 323–332. doi: 10.1037/h0058944

Bromberg-Martin, E. S., Matsumoto, M., and Hikosaka, O. (2010). Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron* 68, 815–834. doi: 10.1016/j.neuron.2010.11.022

Gross, J., Woelbert, E., Zimmermann, J., Okamoto-Barth, S., Riedl, A., and Goebel, R. (2014). Value signals in the prefrontal cortex predict individual preferences across reward categories. *J. Neurosci.* 34, 7580–7586. doi: 10.1523/JNEUROSCI.5082-13.2014

Haber, S. N., and Knutson, B. (2010). The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology* 35, 4–26. doi: 10.1038/npp.2009.129

Hoffeld, D. R., Kendall, S. B., Thompson, R. F., and Brogden, W. J. (1960). Effect of amount of preconditioning training upon the magnitude of sensory preconditioning. *J. Exp. Psychol.* 59, 198–204. doi: 10.1037/h0048857

Howard, J. D., Gottfried, J. A., Tobler, P. N., and Kahnt, T. (2015). Identity-specific coding of future rewards in the human orbitofrontal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 112, 5195–5200. doi: 10.1073/pnas.1503550112

Howard, J. D., and Kahnt, T. (2017). Identity-Specific reward representations in orbitofrontal cortex are modulated by selective devaluation. *J. Neurosci.* 37, 2627–2638. doi: 10.1523/JNEUROSCI.3473-16.2017

Howard, J. D., and Kahnt, T. (2018). Identity prediction errors in the human midbrain update reward-identity expectations in the orbitofrontal cortex. *Nat. Commun.* 9, 1611. doi: 10.1038/s41467-018-04055-5

Howard, J. D., and Kahnt, T. (2021). To be specific: the role of orbitofrontal cortex in signaling reward identity. *Behav. Neurosci.* 135, 210–217. doi: 10.1037/bne0000455

Howard, J. D., Reynolds, R., Smith, D. E., Voss, J. L., Schoenbaum, G., and Kahnt, T. (2020). Targeted stimulation of human orbitofrontal networks disrupts outcome-guided behavior. *Curr. Biol.* 30, 490–498.e4. doi: 10.1016/j.cub.2019.12.007

Huang, Y. Z., Edwards, M. J., Rounis, E., Bhatia, K. P., and Rothwell, J. C. (2005). Theta burst stimulation of the human motor cortex. *Neuron* 45, 201–206. doi: 10.1016/j.neuron.2004.12.033

Jones, J. L., Esber, G. R., Mcdannald, M. A., Gruber, A. J., Hernandez, A., Mirenzi, A., et al. (2012). Orbitofrontal cortex supports behavior and learning using inferred but not cached values. *Science* 338, 953–956. doi: 10.1126/science.1227489

Kahnt, T. (2018). A decade of decoding reward-related fMRI signals and where we go from here. *Neuroimage* 180, 324–333. doi: 10.1016/j.neuroimage.2017.03.067

Levy, D. J., and Glimcher, P. W. (2011). Comparing apples oranges: using reward-specific reward-general subjective value representation in the brain. *J. Neurosci.* 31, 14693–14707. doi: 10.1523/JNEUROSCI.2218-11.2011

Mcnamee, D., Rangel, A., and O'doherty, J. P. (2013). Category-dependent and category-independent goal-value codes in human ventromedial prefrontal cortex. *Nat. Neurosci.* 16, 479–485. doi: 10.1038/nn.3337

Murray, E. A., Moylan, E. J., Saleem, K. S., Basile, B. M., and Turchi, J. (2015). Specialized areas for value updating and goal selection in the primate orbitofrontal cortex. *Elife* 4, e009. doi: 10.7554/eLife.11695.009

Padoa-Schioppa, C. (2011). Neurobiology of economic choice: a good-based model. *Annu. Rev. Neurosci.* 34, 333–359. doi: 10.1146/annurev-neuro-061010-113648

Reber, J., Feinstein, J. S., O'doherty, J. P., Liljeholm, M., Adolphs, R., and Tranel, D. (2017). Selective impairment of goal-directed decision-making following lesions to the human ventromedial prefrontal cortex. *Brain* 140, 1743–1756. doi: 10.1093/brain/awx105

Rudebeck, P. H., Saunders, R. C., Prescott, A. T., Chau, L. S., and Murray, E. A. (2013). Prefrontal mechanisms of behavioral flexibility, emotion regulation and value updating. *Nat Neurosci* 16, 1140–1145. doi: 10.1038/nn.3440

Sadacca, B. F., Jones, J. L., and Schoenbaum, G. (2016). Midbrain dopamine neurons compute inferred and cached value prediction errors in a common framework. *Elife* 5, e010. doi: 10.7554/eLife.13665.010

Sadacca, B. F., Wied, H. M., Lopatina, N., Saini, G. K., Nemirovsky, D., and Schoenbaum, G. (2018). Orbitofrontal neurons signal sensory associations underlying model-based inference in a sensory preconditioning task. *Elife* 7, e30373. doi: 10.7554/eLife.30373

Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599. doi: 10.1126/science.275.5306.1593

Schwartenbeck, P., Fitzgerald, T. H. B., and Dolan, R. (2016). Neural signals encoding shifts in beliefs. *Neuroimage* 125, 578–586. doi: 10.1016/j.neuroimage.2015.10.067

Sharpe, M. J., Batchelor, H. M., and Schoenbaum, G. (2017). Preconditioned cues have no value. *Elife* 6, e006. doi: 10.7554/eLife.28362.006

Stalnaker, T. A., Howard, J. D., Takahashi, Y. K., Gershman, S. J., Kahnt, T., and Schoenbaum, G. (2019). Dopamine neuron ensembles signal the content of sensory prediction errors. *Elife* 8, e010. doi: 10.7554/eLife.49315.010

Suarez, J. A., Howard, J. D., Schoenbaum, G., and Kahnt, T. (2019). Sensory prediction errors in the human midbrain signal identity violations independent of perceptual distance. *Elife* 8, e032. doi: 10.7554/eLife.43962.032

Takahashi, Y. K., Batchelor, H. M., Liu, B., Khanna, A., Morales, M., and Schoenbaum, G. (2017). Dopamine neurons respond to errors in the prediction of sensory features of expected rewards. *Neuron* 95, 1395–1405.e3. doi: 10.1016/j.neuron.2017.08.025

Wang, F., Howard, J. D., Voss, J. L., Schoenbaum, G., and Kahnt, T. (2020a). Targeted stimulation of an orbitofrontal network disrupts decisions based on inferred, not experienced outcomes. *J. Neurosci.* 40, 8726–8733. doi: 10.1523/JNEUROSCI.1680-20.2020

Wang, F., Schoenbaum, G., and Kahnt, T. (2020b). Interactions between human orbitofrontal cortex and hippocampus support model-based inference. *PLoS Biol.* 18, e3000578. doi: 10.1371/journal.pbio.3000578

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Kahnt. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Behavioral Neuroscience

TYPE Opinion
PUBLISHED 18 October 2022
DOI 10.3389/fnbeh.2022.942561

Check for updates

OPEN ACCESS

EDITED BY
Rüdiger J. Seitz,
Heinrich Heine University of Düsseldorf, Germany

REVIEWED BY
Raimund Kleiser,
Kepler University Hospital GmbH, Austria

*CORRESPONDENCE
Philipp Kranabitl
philipp.kranabitl@tugraz.at

SPECIALTY SECTION
This article was submitted to Emotion Regulation and Processing, a section of the journal Frontiers in Behavioral Neuroscience

RECEIVED 12 May 2022
ACCEPTED 06 September 2022
PUBLISHED 18 October 2022

CITATION
Kranabitl P and Faustmann C (2022) The SMH approach as support for decision-making in a technical context.
Front. Behav. Neurosci. 16:942561.
doi: 10.3389/fnbeh.2022.942561

COPYRIGHT
© 2022 Kranabitl and Faustmann. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

The SMH approach as support for decision-making in a technical context

Philipp Kranabitl* and Clemens Faustmann

Institute of Machine Components and Methods of Development, Graz University of Technology, Graz, Austria

KEYWORDS

decision making, bias, decision traps, model based development approach, systems thinking (ST), method

Introduction

In society there is this widely common assumption, that engineers when developing a mechatronic product only make quantifiable and rational decisions. There are many details which lead to a successful product development, but the most important factor considering all aspects is sound decision-making. Eriksson even suggested, that product development can be viewed as a decision production system (Eriksson, 2009). Based on many failed engineering projects, the assumption that engineers mostly make rational decisions must be questioned. How to support decision-making is therefore a central topic in product development. In this paper the so called SMH Approach (S: system-thinking, M: model based, H: human factor) (Kranabitl et al., 2021) is presented. This short paper does not claim to address all relevant aspects of this far-reaching Research Topic in detail but aims to give an overview of the core steps which improve the decision-making capabilities of engineers based on a methodological concept in order to maximize the chances of success for product development. To better understand the human factor, creditions are integrated as a so far neglected human capacity which is rooted in brain function (Angel and Seitz, 2016).

In literature, many approaches exist, that have the purpose to support decision-makers. Many of them are summarized under the term "decision support systems," defined by Little as model-based set of procedures for processing data and judgments to assist a manager in his decision-making (Little, 2004).

When developing a powertrain, a usual method which is applied after the design phase is the FMEA (failure mode and effect analysis). In this method, engineers from different disciplines (mechanical, electric/electronics, and software) work together to define and assess possible weak spots of the design. Engineers have to define possible failures and rate them on a scale from one to ten in three categories (severity, occurrence, and detection). This is considered one of the key methods in product development since design changes and expensive verification activities are based on the outcome of this method. The comparison of powertrain projects showed, that engineers made different assessment to similar failure modes on different occasions. Therefore, to propose a concept for decision support, one has to consider human aspects of decision-making processes. Kahneman (2003) distinguishes between two thinking systems. First, he describes system 1 that summarizes intuitive thinking which is fast, automatic, effortless, emotional and implicit. Secondly, he describes system 2 thinking, that is seen as slower

but conscious, effortful, explicit, and logical way of thinking. In an ideal approach the decision-maker should rely on system 2 thinking, while considering that system 1 thinking happens unconsciously and can only be influenced by knowing the circumstances of the decision, such as the situation, the emotions regarding the subjects etc.

An approach to support decisions therefore has to consider the following aspects:

- Structured illustration of the decision context and possible affected aspects.
- Traceability of the information that is the base for the decision to be made.
- Emphasize the impact of the human factor on the decision output.

The developed decision-support concept, the so-called SMH approach (Kranabitl et al., 2021), considers these aspects and reflects them in its methodical steps.

The SMH approach

The situation for companies in the automotive domain is defined by vast competition, cost pressure, legislative regulations, the emergence of new technologies, and changing customer expectations. In consequence, the future of companies in this industry heavily depends on making the right decisions in both a technical and strategic sense. To deepen the understanding of the presented theory in this paper, an example will be discussed with every step of the SMH approach. On what powertrain variant a company should focus on and consequently invest R&D budget over the next years to maintain profitable on a very competitive market is discussed. A wrong decision on questions like this, could cost companies large amounts of market share and therefore have to be considered very thorough.

Figure 1 shows the decision situation with options to choose from (powertrain variants), the three steps of the SMH approach (which are discussed in detail in the next subsections) and the option chosen based on the evaluation of the SMH approach.

Systems thinking to define decision context

To define the context of a decision, including the situation it is made within, the stakeholders of the decision as well as the influences on society environment and more, systems thinking is necessary. As a key principle of systems engineering, systems thinking is a philosophy and state of mind to think beyond the system under development (Haberfellner et al., 2015). The decision on which powertrain to invest R&D resources isn't only a technical question. Political decisions, customer perception, environmental situations, infrastructure development and many more aspect also play a role in whether the developed product will be a success on the market or not. Therefore, the aim of this first step of the SMH approach is to widen the view of the decision maker in order to consider all relevant aspects for this decision.

To visualize the results of this first step, a semantic network including various aspects that are linked to this decision is considered. A semantic network basically consists of nodes (objects) and links (relations) between them (Bajzek et al., 2021). This developed network can be reused and adapted in future projects and illustrates the considered aspects of possible implications of a decision.

Information base in form of models

As described in the first step, a decision-making process takes several aspects into account. To describe these different aspects, many models are required. In this second step of the approach, these models are collected, adapted, generated. Models in this matter reach from a technical model like a structural or functional model of a subsystem of the powertrain to models which aim to represent the preferences of future customer or political situations. Of course, models of future customer or political situations can only be modeled with a high degree of uncertainty. Yet these models are of high importance. Models with high and known uncertainty (the decision-maker should be aware of the model accuracy) should be embedded into the model landscape and updated when necessary instead of being not considered due to lack of significance. If a model with high uncertainty turns out to be wrong after some time, it needs to be updated and the changes to the whole decision situation have to be evaluated.

In the SMH approach all models are structured with the concept of a three-dimensional model cube to maintain an overview. This cube represents the sum of the considered models and structures these models in three dimensions: discipline, technical domain and level (Hick et al., 2019).

Consideration of the human factor

Influences on decision can be understood by considering the way human brains work (Hammond et al., 2015). In literature, some of these influences are summarized in decision traps and biases by scientists such as Hammond et al. (2015), Korhonen and Wallenius (2020), or Kahneman et al. (2021).

The SMH approach includes human influences on a decision by forcing its applicant to consider decision traps and credition aspects (Hick et al., 2021). While the human influences are not yet quantified in the SMH approach, it provides a framework for decision-makers to consider the anchoring trap, status-quo trap, sunk cost trap, confirming evidence trap, simplicity

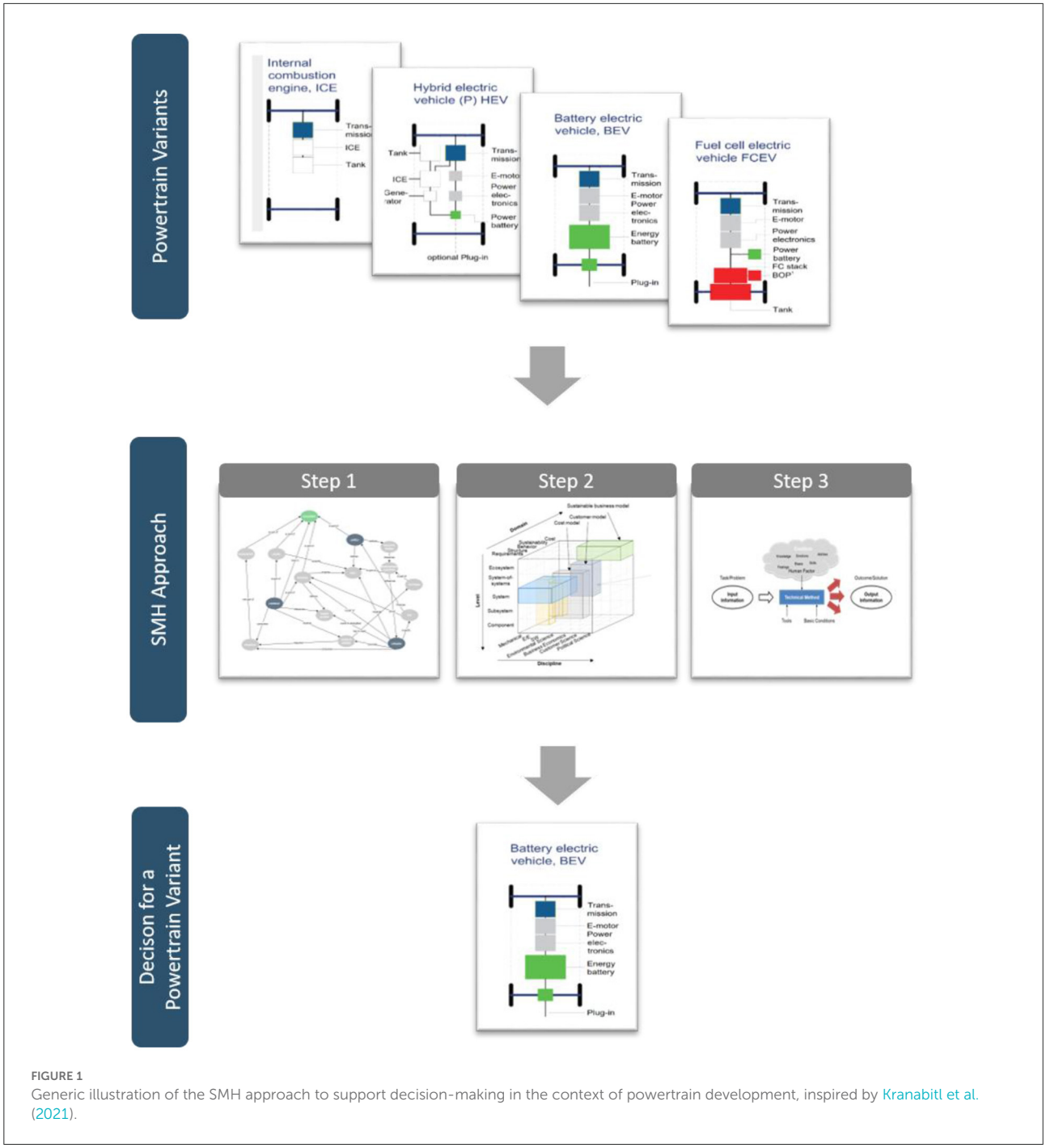


FIGURE 1
Generic illustration of the SMH approach to support decision-making in the context of powertrain development, inspired by Kranabitl et al. (2021).

trap and credition aspects in a structured way. As discussed in step two, models which have a low level of accuracy due to high uncertainty need to be interpreted as such by the decision maker. Considering the confirming evidence bias for this matter, a model which is of low accuracy but confirms the decision-makers initial believe, may be not considered as the vague information it represents but as way more accurate. Misjudgments due to flaws in human judgement like these are a threat to good decision-making and need to be reduced by following a proven process. This process can consist of a few simple questions which force the decision-maker to question if the derived belief is influenced by one of those biases. It can also dictate different people who have to confirm those beliefs in order to continue. The C-E-C (cognitions, emotions, creditions) triangle by Angel which states that forming a belief or believing is not possible without an associated emotion and cognition,

forms the basis for the considered aspects in this step (Angel, 2017).

Discussion

This short article does not aim to explain all steps of the SMH approach in detail but give an overview of core tasks and objectives. Further research is required to develop concepts and eventually applicable methods for persons in charge that have to make decisions. The chosen example of a powertrain variant selection illustrates that such a decision has many possible implications on the company, the society and the environmental system. A decision that appears to be solely of technical nature at first sight, such as the described example, has huge impact on the company's future economic success which leads to investment in research for new technologies or more employees. This directly influences society by providing work for people and by taxes as a result of sales and profits. Furthermore, the decision for a technology affects the environment because of required resources, CO_2 equivalent emissions, and many more aspects.

The SMH approach is a concept to support responsible persons in taking decisions with a technical context. As this approach is more described on sja theoretical basis, future work regarding its implementation has to make sure that the following prerequisite are considered and fulfilled:

- Quantified parameters based on the models and the semantic network to better illustrate the possible options.
- Implementation in form of a method that is simplified to a certain extent, to ensure quick application also for decision-makers that are non-researchers in decision-making theory.
- Providing traceability of taken decisions and the information they are based on.

The SMH approach describes steps to extensively prepare the base or input for a decision by analyzing the situation, by relying on models as main source of information, and by considering the human factors in decision-making. It does not describe how the ideal trade-off between several factors is identified, such as in classical multicriteria decision models. The decision-maker is still challenged to draw conclusions out of the developed decision input in form of the semantic network and models. In the future this approach has to be enhanced and implemented as an IT solution to provide benefits for decision-makers in the sense of an expert system (Butler et al., 1997).

A further interesting addition to the proposed approach is to consider uncertainties in a quantified way and to apply statistics (Pfeifer and Lüthi, 1987). Especially for models which describe a future aspect, such as a model that describes the customer behavior when using the vehicle, a degree of uncertainty has to be considered. E.g., the customer might use the car in 10 years as a shared vehicle with other people rather than using it alone for weekend trips.

Author contributions

CF was in a key role for bringing the core idea to paper. All authors contributed to the article and approved the submitted version.

Funding

CF and PK declare that this study received funding from Siemens Healthineers. The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article, or the decision to submit it for publication.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Angel, H.-F. (2017). "Credition: from the question of belief to the question of believing," in *Processes of Believing: The Acquisition, Maintenance, and Change in Creditions*, eds H. F. Angel, Oviedo, L., Paloutzian, R. F., Runehov, A. L. C., and Seitz, R. J. (Cham: Springer), 17–36.

Angel, H.-F., and Seitz, R. J. (2016). Process of believing as fundamental brain function: the concept of credition. *SFU Res Bull.* 8, 1–12. doi: 10.15135/2016.4.1.1-20

Bajzek, M., Fritz, J., and Hick, H. (2021). "Systems engineering principles," in *Systems Engineering for Automotive Powertrain Development*, eds H. Hick, K. Küpper, H. Sorger (Cham: Springer Nature).

Butler, J., Jia, J., and Dyer, J. (1997). Simulation techniques for the sensitivity analysis of multi-criteria decision models. *Eur. J. Oper. Res.* 103, 531–546. doi: 10.1016/S0377-2217(96)00307-4

Eriksson. (2009). *Collaborative Product Development: A Collaborative Decision-Making Approach*. Mälardalen University Press Licentiate Theses.

Haberfellner, R., de Weck, O., Fricke, E., and Vössner, S. (2015). *Systems Engineering - Grundlagen und Anwendung*. Zürich: Orell Füssli; ed. 13.

Hammond, J., Keeney, R., and Raiffa, H. (2015). *Smart Choices: A Practical Guide to Making Better Decisions*. Brighton: HBR Press.

Hick, H., Angel, H.-F., Philipp, K., and Jolana, W.-S. (2021). "Decision-making and the influence of the human factor," in *Systems Engineering for Automotive Powertrain Development*, eds H. Hick, K. Küpper, H. Sorger (Cham: Springer Nature), 355–380.

Hick, H., Bajzek, M., and Faustmann, C. (2019). Definition of a system model for model-based development. *SN Appl. Sci*. 1. doi: 10.1007/s42452-019-1069-0

Kahneman, D. (2003). A perspective on judgment and choice: mapping bounded rationality. *Am. Psychol*. 58, 697–720. doi: 10.1037/0003-066X.58.9.697

Kahneman, D., Sibony, O., and Sunstein, C. R. (2021). *Noise—A Flaw in Human Judgment, 1st Edn*. New York, NY: Little Brown Spark.

Korhonen, P., and Wallenius, J. (2020). *Making Better Decisions*. Cham: Springer Nature. doi: 10.1201/b18823-11

Kranabitl, P., Faustmann, C., and Hick, H. (2021). Decision-making for sustainable technical applications with the SMH approach. *Sustainability* 13, 8702. doi: 10.3390/su13168702

Little, J. (2004). Models and managers: the concept of a decision calculus. *Manag. Sci*. 50, 1841–1853. doi: 10.1287/mnsc.1040.0267

Pfeifer, R., and Lüthi, H. (1987). "Decision support systems and expert systems: a complementary relationship? In expert systems and artificial intelligence in decision support systems," in *Proceedings of the Second Mini Euroconference*, eds H. Sol, C. Takkenberg, P. Robbe (Lutheren: Springer-Science+ Business Media), 41–51.

frontiers | Frontiers in Psychology

OPINION
published: 11 April 2022
doi: 10.3389/fpsyg.2022.868903

The Synergies Between Understanding Belief Formation and Artificial Intelligence

*Sara Lumbreras**

Institute for Research in Technology, Universidad Pontificia Comillas, Madrid, Spain

Understanding artificial intelligence (AI) and belief formation have interesting bidirectional synergies. From explaining the logical derivation of beliefs and their internal consistency, to giving a quantitative account of mightiness, AI still has plenty of unexploited metaphors that can illuminate belief formation. In addition, acknowledging that AI should integrate itself with our belief processes (mainly, the capacity to reflect, rationalize, and communicate that is allowed by semantic coding) makes it possible to focus on more promising lines such as Interpretable Machine Learning.

Keywords: artificial intelligence, belief, machine bias, complexity, reinforcement learning

OPEN ACCESS

Edited by:
Rüdiger J. Seitz,
Heinrich Heine University of Düsseldorf, Germany

Reviewed by:
Hans Van Eyghen,
VU Amsterdam, Netherlands

***Correspondence:**
Sara Lumbreras
slumbreras@comillas.edu

Specialty section:
This article was submitted to Theoretical and Philosophical Psychology, a section of the journal Frontiers in Psychology

Received: *03 February 2022*
Accepted: *02 March 2022*
Published: *11 April 2022*

Citation:
Lumbreras S (2022) The Synergies Between Understanding Belief Formation and Artificial Intelligence. Front. Psychol. 13:868903. doi: 10.3389/fpsyg.2022.868903

INTRODUCTION

The research program "Credition" has provided a solid framework to understand the phenomenon of believing and belief formation.

"*Credition, the processes of believing, are fundamental brain functions that enable a non-human animal or human being to trust his/her inner probabilistic representation(s). Credition is based on neural processes, including perception and valuation of objects and events in the physical and social environment in secular as well as religious contexts. By predictive coding, credition guides one's actions and behaviors through reciprocating feedback involving learning* (Angel et al., 2017)."

Artificial Intelligence (AI) can be understood as a sequence of algorithms (that is, the mechanical application of some predefined steps) that is applied to a set of data and that generates a probabilistic representation of these data with the aim of making predictions, inferring a consequence or selecting the best possible option. AI can be understood as a machine that supports the formation and valuation of beliefs in the human and can be understood metaphorically as a belief-machine itself.

Artificial intelligence encompasses techniques such as clustering or pattern recognition. However, this paper focuses primarily on *Reinforcement Learning* (RL), where intelligent agents take actions in an environment to maximize a reward and punish mistakes. In RL, there are steps that are metaphorically described as perception (collecting new data), decision (selecting the optimal action), valuation (evaluating the outcome of a decision), or learning (the successive improvements in valuation obtained through repeated cycles of decision and valuation). For instance, a RL system could learn to play chess through cycles of selecting a move and valuing the possible positions.

For this reason, it is extremely interesting to examine AI from the lens of the credition process: understanding how AI works can give us insights to inform our hypothesis about the workings of credition. In parallel, acknowledging that AI should support belief formation helps to design it better and make this support as effective as possible.

SYNERGIES FROM UNDERSTANDING ARTIFICIAL INTELLIGENCE TO UNDERSTANDING BELIEF FORMATION

Figure 1, taken from Seitz et al. (in press), presents a schematic representation of how a RL process can be understood in the context of credition, displaying the different levels of memory that are at play (Rolls, 2000). The inner probabilistic representation (which we refer to as "belief") is built around the data received (perception) and is also used to value future actions (valuation). This representation is used to select the preferred course of action and predict its outcome. When new data are collected about the outcome (the *prediction error*), the probabilistic representation is updated through a new valuation in the process of RL. It is important to note that, according to this model, the credition process happens in a spontaneous and automatic manner, below the level of awareness.

For instance, one particularly interesting insight from the Creditions model and RL is the interpretation of the balance of *exploration* (which, in this context, we will understand the examination of new decisions or beliefs not tested before) vs. *exploitation* (the use of the existing belief system to make a decision in an efficient manner). This balance, which is key to the performance of RL algorithms, can be seen as an essential feature of belief formation and update that depends greatly on the psychological characteristics of each individual. In addition, In RL, each new data point is integrated into the probabilistic representation of the world. The specific manner in which this is done can be represented as an error minimization strategy. This model could be tested in experimental settings to improve our current understanding of how beliefs are formed and updated.

In addition, there also are interesting insights outside RL that can help us improve our understanding of the mechanisms of credition.

Illuminating Metaphors From Artificial Intelligence Outside Reinforcement Learning

Reinforcement learning does not capture all the complexities of belief formation and update. Importantly, RL takes a blind approach to the inner probabilistic representation and does not necessarily impose any internal consistency to beliefs. However, we do know that new beliefs are more easily accepted when they are consistent with prior ones (Fryer et al., 2019) or that cognitive dissonance is an unpleasant experience (Cooper, 2007). We also know that some beliefs are derived logically from others. This makes it beneficial to resort also to some AI tools that have an emphasis in this logical consistency, or on the logical derivation of consequences. These are inference engines and Bayesian networks.

Inference Engines and the Logical Derivation of Belief

Inference engines are tools that apply predefined logical rules to a knowledge database (also previously defined) to derive new facts from already known ones. We refer the reader to Colmerauer (1990) for a good introduction to Prolog, one of the first and most widely used inference engines which name is derived from the expression "Propositional Logic," which is the basis of its first version (Prolog I), which was later upgraded to include first-order logic (quantification) or fuzzy logic (Zadeh, 1988) (which allows for intermediate states between true and false). Inference engines work in a similar manner to generic theorem provers such as the well-known *Isabelle* (Paulson, 1994).

Inference engines apply the logical rules first to the facts contained in the knowledge database to derive an initial set of consequences. This is performed by examining every potential pair for a possible conclusion. Then, both the initial facts and the newly obtained consequences are combined again to generate a second round of consequences.

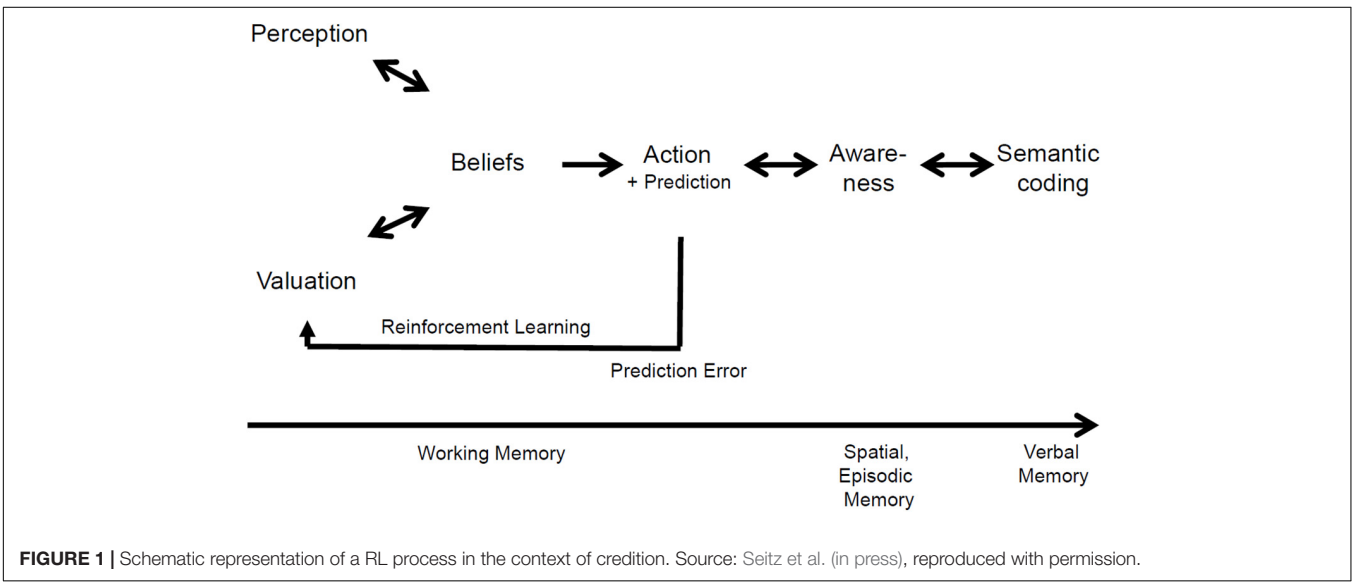


FIGURE 1 | Schematic representation of a RL process in the context of credition. Source: Seitz et al. (in press), reproduced with permission.

This process is repeated iteratively until no more new facts can be derived.

At some level, human beings also scan new knowledge for possible consequences using inference rules (Rodriguez et al., 2016). We could even understand that the more iterations of the process need to be performed in order to find a consequence, the more obscured it will be. It should be useful to incorporate a metaphor of the inference engine to the creditions model to understand how logical consequences are derived from new beliefs at the step of the inner probabilistic construction.

Bayesian Networks and the Internal Consistency of Belief

A Bayesian Network is a probabilistic graphical model that represents a set of variables and their conditional dependencies *via* a directed acyclic graph. This means that variables are connected always with a direction (there is one cause and one consequence, without any possible circularity). Bayesian networks are a very interesting tool for understanding the contributing factor for an event. For example, it can represent the probabilistic relationships between diseases and their symptoms, so we could calculate the specific probabilities that an observed symptom is due to a given disease. We refer the reader to Chen and Pollino (2012) for a good tutorial on this topic. We know that beliefs are not held in isolation but could rather be understood as a network (Friedkin et al., 2016). For this reason, including the remarkable understanding of networks and their relationships that AI has created with Bayesian Networks can be extremely interesting.

Complexity Theory to Understand the Global Properties of Belief Systems

A complex system is a system composed of many elements in interaction. We can find complex systems in contexts as different as the global climate, social organizations or the metabolism within a cell (Mitchell, 2009). Very importantly, the algorithms that support the developments of AI are also complex systems.

Complex systems have distinct properties which are shared by belief systems, which include being goal-oriented, open to receiving information from the exterior, spontaneous order (with hierarchies and context appearing), adaptation, being difficult to predict, experiencing non-linear phenomena (for instance, it is much more difficult to change a belief than to form it). I refer the reader to my paper (Lumbreras and Oviedo, 2020) for a more detailed account of these properties and their implications.

Understanding Mightiness in a Quantitative Manner

In addition to its object, a belief can be characterized by its certainty, which can be very aptly represented by the certainty level in fuzzy logic that is embedded in Bayesian networks. Moreover, there is also a second qualifier: mightiness, which refers to the intensity that is attributed to this emotion. AI can also help us to understand this mightiness at a quantitative level. There is a property of the variables involved in a prediction model that is called *importance*. The importance of a variable is a measure of how much this variable affects the final decision. Arguably, we feel more intensely the beliefs that more profoundly affect our own identity and actions.

DISCUSSION: FROM UNDERSTANDING BELIEF TO CREATING BETTER ARTIFICIAL INTELLIGENCE

As explained in the section "Introduction," the synergies between understanding AI and understanding belief are bidirectional, so there are positive outcomes that can be expected if we introduce what we know about belief formation into the way we design and use AI.

There are two main issues that plague the applications of AI: overfitting and machine bias. We need to remember that AI extracts patterns from the data that it receives for training, so it deeply depends on these data.

In overfitting, the data provided to the machine is not enough to be able to generalize. Much like a student that, instead of understanding, learns examples by heart, the algorithm fails when a new situation is considered.

In the same way, it is possible that the data we present to the algorithm does not represent reality fairly. For instance, it has been well documented that some algorithms disfavor African Americans, for instance when calculating their probability of recidivism in crime with the objective of deciding whether to grant them parole (Hajian et al., 2016). This was due to the database that was used for training containing a higher proportion of African American recidivists.

The problem with overfitting and machine bias is that they are not easy to detect. Most applications of AI are designed as *black boxes,* so we only have access to their specific predictions for every case but not to any reasoning behind them. Without detailed analyses, for instance, it is not possible to determine that the prediction is based on race, age, sex, or any other discriminatory variable. This means that when black-box AI is used in a high-stake decision (such as granting parole), this can have disastrous consequences.

As I have presented, black-box AI only shares predictions, and this can have disastrous consequences in high-stake problems. However, there is an emerging field within AI, *Interpretable ML*, which does allow for the understanding of models and their dynamics. This makes it possible avoid the problems derived from overfitting and machine bias (Molnar, 2020). The basic idea behind Interpretable ML is that, for many problems, it is possible to create AI that is so simple that can be expressed in rules understood by a human, and yet result in accurate predictions. These models would be the opposite to black boxes, they are transparent decision rules that can be understood and discussed.

For instance, Rudin (2019) developed an alternative to the parole algorithm that was based only on prior violent crimes and

age. This shows that it is possible to create AI that better fits the way we form beliefs and is a more efficient support for our decision making.

CONCLUSION

Understanding AI and understanding belief formation have interesting bidirectional synergies. From explaining the logical derivation of beliefs and their internal consistency, to giving a quantitative account of mightiness, AI still has plenty of metaphors to illuminate belief. Potentially, these can be used to simulate belief systems and arrive to testable predictions.

In addition, acknowledging what our belief processes have that AI lacks (mainly, the capacity to reflect, rationalize and communicate that is allowed by semantic coding) makes it possible for us to focus on creating AI that can better support decisions such as Interpretable ML.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

REFERENCES

Angel, H., Oviedo, L., Paloutzian, R. F., Runehov, A. L., and Seitz, R. J. (2017). *Processes of believing: the acquisition, maintenance, and change in creditions*. Berlin: Springer.

Chen, S. H., and Pollino, C. A. (2012). Good practice in bayesian network modelling. *Environ. Model. Softw.* 37, 134–145. doi: 10.1016/j.envsoft.2012.03.012

Colmerauer, A. (1990). *An introduction to prolog III. Paper presented at the Computational Logic*, 37-79. New York, NY: Association for Computing Machinery.

Cooper, J. (2007). *Cognitive dissonance: 50 years of a classic theory*. Thousand Oaks: Sage

Friedkin, N. E., Proskurnikov, A. V., Tempo, R., and Parsegov, S. E. (2016). Network science on belief system dynamics under logic constraints. *Science* 354, 321–326. doi: 10.1126/science.aag2624

Fryer, R. G. Jr., Harms, P., and Jackson, M. O. (2019). Updating beliefs when evidence is open to interpretation: implications for bias and polarization. *J. Eur. Econ. Assoc.* 17, 1470–1501. doi: 10.1093/jeea/jvy025

Hajian, S., Bonchi, F., and Castillo, C. (2016). "Algorithmic bias: from discrimination discovery to fairness-aware data mining," in *Paper presented at the Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, (New York: ACM, Inc), 2125–2126.

Lumbreras, S., and Oviedo, L. (2020). Belief networks as complex systems. *LIMINA-Grazer Theologische Perspektiven* 3, 92–108.

Mitchell, M. (2009). *Complexity: a guided tour*. Oxford: Oxford University Press.

Molnar, C. (2020). *Interpretable machine learning*. Abu Dhabi: Lulu. com.

Paulson, L. C. (1994). *Isabelle: a generic theorem prover*. Berlin: Springer Science & Business Media.

Rodriguez, N., Bollen, J., and Ahn, Y. Y. (2016). Collective dynamics of belief evolution under cognitive coherence and social conformity. *PLoS One* 11:e0165910. doi: 10.1371/journal.pone.0165910

Rolls, E. T. (2000). Memory systems in the brain. *Annu. Rev. Psychol.* 51, 599–630. doi: 10.1146/annurev.psych.51.1.599

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat. Mach. Intell.* 1, 206–215. doi: 10.1038/s42256-019-0048-x

Seitz, R. J., Angel, H., and Paloutzian, R. F. (in press). Bridging the gap between believing and memory function. *Eur. J. Psychol.* doi: 10.23668/psycharchives.5421

Zadeh, L. A. (1988). Fuzzy logic. *Computer* 21, 83–93.

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Lumbreras. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Behavioral Neuroscience

OPINION
published: 22 June 2022
doi: 10.3389/fnbeh.2022.928297

Evolving Belief, Evolving Minds: Evolutionary Insights Into the Development and Functioning of Human Society

*Agustín Fuentes**

Department of Anthropology, Princeton University, Princeton, NJ, United States

Keywords: evolution, neurobiology, human mind, genus *Homo*, belief

OPEN ACCESS

Edited by:
Rüdiger J. Seitz, Heinrich Heine University of Düsseldorf, Germany

Reviewed by:
Martin Porr, University of Western Australia, Australia

***Correspondence:**
Agustín Fuentes
afuentes2@princeton.edu

Specialty section:
This article was submitted to Emotion Regulation and Processing, a section of the journal Frontiers in Behavioral Neuroscience

Received: *25 April 2022*
Accepted: *30 May 2022*
Published: *22 June 2022*

Citation:
Fuentes A (2022) Evolving Belief, Evolving Minds: Evolutionary Insights Into the Development and Functioning of Human Society. Front. Behav. Neurosci. 16:928297. doi: 10.3389/fnbeh.2022.928297

Belief involves the ability to think beyond what is here and now and develop mental representations in order to see and feel and know something not immediately present to the senses, and invest in that something so that it becomes one's reality (Fuentes, 2019). Belief involves mental activity constituted by neural circuits in the brain (Boyer, 2003), but it is more than that. Belief involves the human ability to draw on cognitive and social resources, histories and experiences, and combine them with imagination to produce neurobiological, physiological, mental and social experience. Belief is a capacity, which may include manifestations of a mental state or attitude involving the appraisal of a proposition, but is not simply this particularly cognitively complex human ability of perceptive and affective information processing (e.g. Seitz and Angel, 2020). Nor is belief solely a property arising from the human capacity for extensive shared agency and shared intentionality (e.g. Tomasello, 2019), although both of those processes form aspects of the human capacity for belief. The capacity for belief enables the human to commit wholly and fully to an idea, a sensation, a concept such that it structures perceptual and experiential processes. Beliefs and belief systems permeate contemporary human neurobiologies, bodies, and ecologies, acting as dynamic agents in evolutionary processes and playing core roles in structuring human societies and the human mind (Stotz, 2010; Downey and Lende, 2012; Han, 2017; Fuentes, 2019; Seitz and Angel, 2020).

Belief is not an 'emergent property', something ephemeral floating above the material reality of being human. It is a central component of the human experience. The ability to believe is part of the human system similar to the way that fingers are part of our arms and hands. Fingers are core aspects of human anatomy, modified over evolutionary time dramatically expanding our options for interacting with the world and each other. In humans, mammalian and then primate limbs were shaped and altered over evolutionary time so that their ends contain structures (prehensile digits and hands with precision grips) expanding the capacities for engagement with, and manipulation of, the world. The capacity for belief is similar: it expands human cognitive, sensory and perceptual dynamics and is critical in the human ability to engage with and shape the world.

In an evolutionary context, beliefs provide for both novel alterations and continued coherence in the human niche. In this brief essay I outline key elements in human evolutionary history that facilitated the emergence of the capacity for belief and suggest that beliefs act as core niche constructive processes in the development of the human mind.

EVOLUTIONARY CONTEXT AND HISTORY

A niche is the structural, temporal, and social dynamic in which a species exists. The niche involves the interfaces between individuals and space, structure, climate, nutrients, and other physical and social factors as a dynamic set of interacting processes (Wake et al., 2009). Over the last two million

years members of the genus *Homo* (humans) underwent significant changes via the emergence of a distinctively human niche. Relative to other hominins, *Homo* underwent specific morphological changes alongside significant behavioral, ecological and cognitive shifts as they forged and were shaped by this human niche (Fuentes, 2015; Marks, 2015; Antón and Kuzawa, 2017; Kissel and Fuentes, 2021). During this time core human patterns emerged including: hyper-cooperation and complex collaboration in social interactions and material technologies; substantially extended childhood development and complex caretaking behavior; intricate and diverse foraging and hunting patterns involving complex technologies, behavior and communication; novel and dynamic material and symbolic cultures eventually resulting in complex cognitive and material meaning-making processes; emergence of exchange networks and increasingly dynamic intergroup relations; and increasingly complex communication and information sharing, eventually resulting in language (Foley, 2016; Fuentes, 2017, 2018; Galway-Witham et al., 2019) (**Figure 1**).

Across the last million years there were many morphologically and behaviorally diverse populations of the genus *Homo* occupying and shaping the human niche, initially across Africa and Eurasia and eventually into Australasia, the Americas and multiple islands across the planet. The taxonomic distinctions between these populations are far from clear. Some argue for multiple species and others for many subspecies, with others suggesting that it is not currently possible to determine the correct number and types of taxa within the genus *Homo* (Schwartz and Tattersall, 2015; Wood and Boyle, 2016). Given the morphological and ecological diversity, and the multiple tool technologies and lifeways evident across this period it is clear that there were many successful ways to navigate the human niche and that they all were intricately connected to, and stemming from, an evolving cognitive capacity setting the stage for the contemporary human mind. Contemporary *Homo sapiens* are inheritors of a diversity of biological and cultural histories facilitated by the dynamics of the human niche (Kissel and Fuentes, 2021).

Current integrative approaches to human evolution emphasize mutual mutability between agents, bodies, collective action, social perceptions, and the roles of experiences, cultures and institutions in structuring human behavior (Fuentes, 2009, 2017; Marks, 2012; Fry, 2013; Kim and Kissel, 2018; Seitz et al., 2018; Sykes, 2020; Kissel and Fuentes, 2021; e.g. DeSilva, 2021). Such complex and multifarious dynamics model interfaces of ecological, behavioral, cultural and cognitive processes as core in the human niche enabling conjectures about the processes at play in the emergence of distinctively human cognition and thought, a "human mind."

In previous work (Fuentes, 2015, 2016, 2017) I've argued for envisioning the human niche as encompassing individual bodies and their evolutionary histories and the patterns and dynamics of interactions within social groups, interactions among/between social groups, and at the community/population levels all within an interactive dynamic with local ecologies (see also Whiten and Erdal, 2012; Foley, 2016). In such a model, evolutionary processes exert pressures at various nodes in the system and responses to those pressures emerge at individual, group, and community levels. The human niche is a dynamic produced by proactive and reactive responses to social and ecological pressures and contexts at various levels creating local and regional ecologies of interactive material, social, cognitive, and historical aspects that flow from one generation to the next; it creates a shared ecology across time and space, the cultural context in which humans evolve (Henrich, 2016; Fuentes, 2017; Laland, 2017; Boyd, 2018). In the development of that human niche the capacity for belief emerged as a significant component creating a dynamic suite of affordances and constraints on human lives facilitated through human cognition, perception, and thought. The evolution of the human niche then, included the emergence of a shared imagination and a suite of distinctive socio-cognitive processes (Whiten and Erdal, 2012; Tomasello, 2014; Fuentes, 2017; Laland, 2017) and a ubiquitous semiotic ecosystem (Deacon, 2016) as central to the context in which humans evolve.

MEANING MAKING, CULTURE, AND CONTEMPORARY HUMAN COGNITION

The environment humans make for themselves is created through their symbol using ability, their capacity for abstraction. The symbols, the ideas, are created in the mind... but the human animal learns not only to create them, but to project them onto the external world, and there transform them into reality. –Montagu (1965), The Human Revolution [1965:2–3]

The patterns and processes of contemporary human cognition and culture, the human perceptual landscape and core facets of human minds, emerged alongside the processes of toolmaking, foraging, caretaking, the control of fire, the creation of symbolic materials, and the ecological expansion of humans across the planet. This ongoing dynamic, the feedback between neural and behavioral plasticity, laid the neurobiological, social, and ecological foundations in human populations for a particularly complex cognition, and for belief (Deacon, 1997; Fuentes, 2019; Tomasello, 2019; Corbey, 2020). The ratcheting up of social and ecological complexity, combined with increased interactions among populations of the genus *Homo*, particularly over the last 200,000 to 500,000 years, created opportunities for the connections and exchanges between groups and populations that enabled shared beliefs, and eventually belief systems, to emerge (Galway-Witham et al., 2019; Kissel and Fuentes, 2021). The last few hundred thousand years offer material evidence for an increase in, and eventual ubiquity of, meaning-making, art and symbol in human populations (Malafouris, 2013; Deacon, 2016; Roberts, 2016; Fuentes, 2017; Sykes, 2020). Across this process humans developed a capacity for imagination and conceptual innovation. These cognitive processes entailed the emergence of two significant patterns. First, the imagining of novel items and/or representations and either making them or altering other things to become them. Such a capacity appears in a limited form in other animals but becomes permanently and ubiquitously part of the human niche by the middle to late Pleistocene. Second and drawing on the first, over the last few hundred thousand years of our history, as part of our intensive communicative and semiotic capacities, humans began creating explanations of

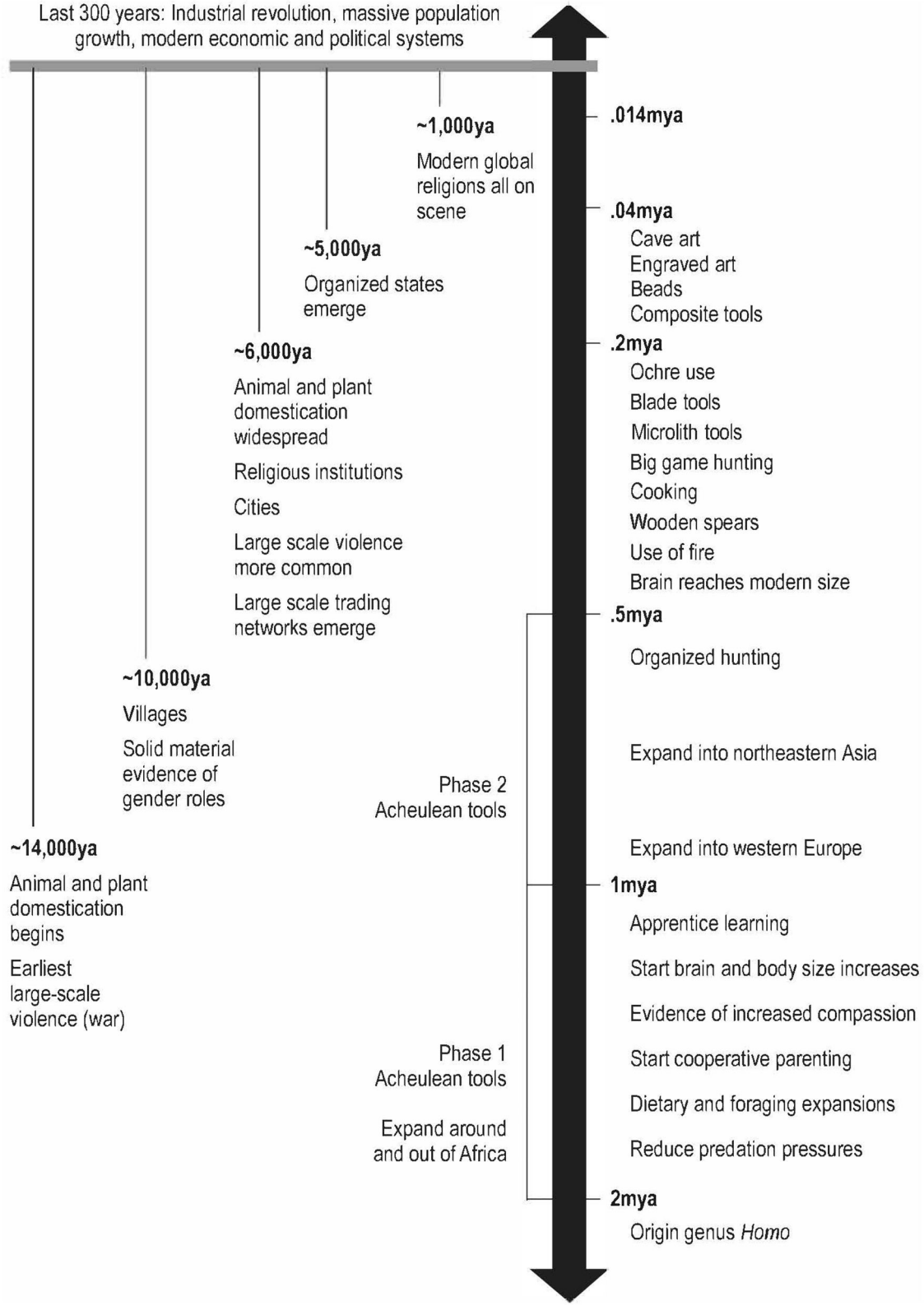


FIGURE 1 | A summary overview of the many key capacities, events and processes in the human niche dated to their first appearances in the fossil and archeological record. Reproduced from Fuentes, 2018, 2019, with permission from University of Chicago Press and Yale University Press.

widely observable phenomena such as death, the behavior of other animals, weather, or the sun and moon. They did not, for example, simply connect clouds, thunder, rain, and floods, they also developed explanations for why these things happen (Deacon, 1997, 2016; Tomasello, 2014; Henrich, 2016; Fuentes, 2017, 2018, 2019; Kissel and Fuentes, 2017). This capacity is what Bloch (2008) refers to in arguing that over evolutionary time humans went from socially complex transactional beings (like most social mammals and other primates) to groups of organisms who exist simultaneously in both transactional and transcendent realities, and who use imagination and belief to reshape themselves and the world around them (Fuentes, 2019).

HUMAN CULTURE/HUMAN MIND

While many organisms have cultures (Whiten, 2021), human culture is demonstrably distinctive. Human culture affects the way that humans do almost everything: fighting, eating, reproducing, innovating, interacting, cooperating, perceiving, making and using technology, expressing ourselves, experiencing emotions, and a host of other cognitive and behavioral processes and events. Culture makes human reason, human being, possible; it forms the central facet of the human niche (Tomasello, 2014; Laland, 2017). Yet individual cultures constrain as much as they enable. Cultures shape social processes and outcomes as well as individuals' development. Cultural contexts, the "webs of significance" that are symbolic meaning, are both materially and perceptually real for the people within them and thus structurally relevant to, and affected by, evolutionary processes and societal processes. When something happens – an action, observation, or experience – our cultural context helps give it meaning, and our participation in that culture enables us to interact with that meaning, making the engagement dynamic and malleable. So, if culture has meaning, then the symbols, ideals, and traditions human participate in come ready-made with relevance and connection to our personal schemata; they make sense to us and shape how we interact with the world. When culture becomes a species' capacity and necessity, as it is for *Homo sapiens*, understanding the mechanisms by which cultural processes evolved, how they function and how such function impacts members and populations of that species itself is of primary interest in any evolutionary narrative of the mind.

For example, a stone tool is not relevant to human evolution simply as the combination of a person altering and using a shaped stone, but rather requires the fact that a person has a set of beliefs, or concepts, of a tool to begin with. The stone object is given shape but also a functional capacity in affecting the world by being transformed from stone to tool, not just through mechanical modification, but also by an understanding about "tool" as a concept. Such assemblages of practical and conceptual processes are a cognitive outcome of evolved capacities in the human niche. A human with the tool concept, and beliefs about the tools themselves, is not constrained by existing tools or materials when novel challenges arise. Rather they can try to innovate and find and modify a stone, or other material, into a novel or altered tool for the job. Likewise, beliefs can shape how social interactions and behavior impact bodies. The contemporary belief of an infant as a fragile (or not) body affects adult handling of infants in ways that influence the maturation processes in a child's motor system, leading to differences in the attainment of landmark events in motor development by working through parental behavior on developmental pathways (Hopkins and Westra, 1990). On a broader populational scale, a shared cultural belief in monotheism can affect social organization and has significant impacts on human reproduction, phenotype, or functioning. It can be linked, for example, to entrenched social inequality such that it makes hierarchy and differential resource distribution more likely to occur, and it increases the likelihood of large state formation or endurance (see Norenzayan, 2013; Henrich, 2020).

Cultural beliefs are important because they fundamentally and reliably change humans' relationships to our environments, the resources at our disposal (e.g., tools, senses, communication), and the conditions of our maturation (the developmental niche), which can have both intra- and intergenerational impact (Seitz et al., 2018; Fuentes, 2019). They are a fundamental part of the niche into which humans born and through which they will interact with the world and other people. Rather than rehashing either side of well-worn debates about the relative importance or contribution of biological and cultural processes, it is evident that the human experience is composed of interacting, co-determining elements of both. And that this process evolved as a central component of the human niche. Human neuroanatomy makes experience material—neural systems adapt through long-term refinement and remodeling, which leads to learning, memory, maturation, which structure perception and affect the creation of beliefs. Through systematic change in the nervous system, and immersion in cultural contexts, humans learn to orchestrate themselves. Cultural concepts and meanings become anatomy (Downey and Lende, 2012). Beliefs infuse human minds, bodies, and ecologies, creating dynamic perceptual and interpretative assemblages that can act either as robust 'enculturalizing' forces in human social systems/socioecologies (our cultures) or disrupt them, facilitating new and/or modified dynamics in perceptual and cultural processes. Therefore belief, and its related cognitive processes and their evolutionary history, matters in assessing human behavior and experience; belief shapes the human mind, past, present, and future.

AUTHOR CONTRIBUTIONS

AF conceived, wrote, and revised all aspects of this article.

ACKNOWLEDGMENTS

My thanks to Rudiger Seitz, Hans-Ferdinand Angel, Raymond Paloutzian, and Ann Taves for organizing the symposium and shepherding this special volume.

REFERENCES

Antón, S. C., and Kuzawa, C. W. (2017). Early homo, plasticity and the extended evolutionary synthesis. *Inter. Focus.* 7, 20170004. doi: 10.1098/rsfs.2017.0004

Bloch, M. (2008). Why religion is nothing special but is central phil. *Trans. R. Soc. B.* 12, 2055–61. doi: 10.1098/rstb.2008.0007

Boyd, R. (2018). *A Different Kind of Animal*. Princeton: Princeton University Press.

Boyer, P. (2003). Religious thought and behaviour as by-products of brain function. *Trends Cogn. Sci.* 7, 119–124. doi: 10.1016/S1364-6613(03)00031-7

Corbey, R. (2020). Baldwin effects in early stone tools. *Evol. Anthropol. Issues News Rev.* 29, 237–244. doi: 10.1002/evan.21864

Deacon, T. (2016). "On human (symbolic) nature: How the word became flesh", in T. Fuchs and C. Tewes (Eds.), *Embodiment in Evolution and Culture*. Tbingen: Mohr Siebeck. p. 129–149.

Deacon, T. W. (1997). *The Symbolic Species: The Co-Evolution of Language and the Brain*. New York, NY: W. W. Norton.

DeSilva (2021). *A Most Interesting Problem. Human evolution 150 years after Darwin's Descent of Man*. Princeton: Princeton University Press. doi: 10.1515/9780691210810

Downey, G., and Lende, D. H. (2012). Evolution and the brain. In D. H. Lende and G. Downey (Eds.), *The Encultured Brain: An Introduction to Neuroanthropology*. Cambridge: MIT Press. p. 103–138.

Foley, R. A. (2016). Mosaic evolution and the pattern of transitions in the hominin lineage. *Philos. Trans. R Soc. B Biol. Sci.* 371, 1–14. doi: 10.1098/rstb.2015.0244

Fry, D. (2013). *War, Peace, and Human Nature*. Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780199858996.001.0001

Fuentes, A. (2009). *Evolution of Human Behavior*. Oxford: Oxford University Press.

Fuentes, A. (2015). Integrative anthropology and the human niche: toward a contemporary approach to human evolution. *Am. Anthropol.* 117, 302–315. doi: 10.1111/aman.12248

Fuentes, A. (2016). The extended evolutionary synthesis, ethnography, and the human niche: toward an integrated anthropology. *Curr. Anthropol.* 57, 13–26. doi: 10.1086/685684

Fuentes, A. (2017). Human niche, human behaviour, human nature. *Interface Focus* 7, 20160136. doi: 10.1098/rsfs.2016.0136

Fuentes, A. (2018). How humans and apes are different, and why it matters. *J. Anthropol. Res.* 74, 151–167. doi: 10.1086/697150

Fuentes, A. (2019). *Why We Believe: Evolution and the Human Way of Being*. Yale University Press/Templeton Press. doi: 10.12987/9780300249255

Galway-Witham, J., Cole, J., and Stringer, C. B. (2019). Aspects of human physical and behavioural evolution during the last 1 million years. *J. Quarter. Sci.* 34, 355–378. doi: 10.1002/jqs.3137

Han, S. (2017). *The Sociocultural Brain*. Oxford UK: Oxford University Press. doi: 10.1093/acprof:oso/9780198743194.001.0001

Henrich, J. (2016). *The Secret of Our Success: How Culture is Driving Human Evolution, Domesticating Our Species, and Making us Smarter*. Princeton: Princeton University Press. doi: 10.1515/9781400873296

Henrich, J. (2020). *The WEIRDest People in the World: How the West Became Psychologically Peculiar and Particularly Prosperous*. New York, NY: Farrar, Straus and Giroux.

Hopkins, B., and Westra, T. (1990). Motor development, maternal expectations, and the role of handling. *Infant Behav. Develop.* 13, 117–122. doi: 10.1016/0163-6383(90)90011-V

Kim, N. C., and Kissel, M. (2018). *Emergent Warfare in Our Evolutionary Past*. New York: Routledge. doi: 10.4324/9781315151021

Kissel, M., and Fuentes, A. (2017). Semiosis in the pleistocene cambridge archaeological. *Journal* 27, 1–16. doi: 10.1017/S095977431700 0014

Kissel, M., and Fuentes, A. (2021). The ripples of modernity: how we can extend paleoanthropology with the extended evolutionary synthesis. *Evolut. Anthropol.* 30, 84–98. doi: 10.1002/evan.21883

Laland, K. N. (2017). *Darwin's Unfinished Symphony: How Culture Made the Human Mind*. Princeton: Princeton University Press. doi: 10.1515/9781400884872

Malafouris, L. (2013). *How Things Shape the Mind: A Theory of Material Engagement*. Cambridge, MA: MIT Press. doi: 10.7551/mitpress/9476.001.0001

Marks, J. (2012). The biological myth of human evolution. *Contemp. Soc. Sci.* 7, 139–157. doi: 10.1080/21582041.2012.691989

Marks, J. (2015). *Tales of the Ex-Apes: How We Think About Human Evolution*. Berkeley: UC Press. doi: 10.1525/california/9780520285811.001.0001

Montagu, A. (1965). *The Human Revolution*. New York: World Publishing Company.

Norenzayan, A. (2013). *Big Gods: How Religion Transformed Cooperation and Conflict*. Princeton, NJ: Princeton University Press. doi: 10.1515/9781400848324

Roberts, P. (2016). We have never been 'behaviourally modern': the implications of Material Engagement Theory and Metaplasticity for understanding the late Pleistocene record of human behaviour. *Quat. Int.* 405, 8–20. doi: 10.1016/j.quaint.2015.03.011

Schwartz, J. H., and Tattersall, I. (2015). Defining the genus *Homo*. *Science*. 349, 931–932. doi: 10.1126/science.aac6182

Seitz, R. J., and Angel, H.-F. (2020). Belief formation—A driving force for brain evolution. *Brain Cogn.* 140, 105548. doi: 10.1016/j.bandc.2020.105548

Seitz, R. J., Paloutzian, R. F., and Angel, H. F. (2018). From believing to belief: a general theoretical model. *J. Cogn. Neurosci.* 9, 1254–1264. doi: 10.1162/jocn_a_01292

Stotz, K. (2010). Human nature and cognitive–developmental niche construction. *Phenom Cogn. Sci.* 9, 483–501. doi: 10.1007/s11097-010-9178-7

Sykes, R. W. (2020). *Kindred: Neanderthal Life, Love, Death and Art*. London: Bloomsbury Sigma. doi: 10.5040/9781472988201

Tomasello, M. (2014). *A Natural History of Human Thinking*. Cambridge: Harvard University Press. doi: 10.4159/9780674726369

Tomasello, M. (2019). *Becoing Human: a Theory of Ontogeny*. Cambridge: Harvard University Press doi: 10.4159/9780674988651

Wake, D. B., Hadly, E. A., and Ackerly, D. D. (2009). Biogeography, changing climates, and niche evolution. *Proc. Nat. Acad. Sci.* 106, 19631–19636. doi: 10.1073/pnas.0911097106

Whiten, A. (2021). The burgeoning reach of animal culture. *Science*. 372, eabe6514. doi: 10.1126/science.abe6514

Whiten, A., and Erdal, D. (2012). The human socio-cognitive niche and its evolutionary origins. *Philos. Trans. R. Soc. B* 367, 2119–2129. doi: 10.1098/rstb.2012.0114

Wood, B., and Boyle, E. (2016). Hominin taxic diversity: fact or fantasy? *Yearbook Phys. Anthropol.* 159, S37–S78. doi: 10.1002/ajpa.22902

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Fuentes. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Psychology

OPINION
published: 29 June 2022
doi: 10.3389/fpsyg.2022.919808

Abstract Concepts, Social Interaction, and Beliefs

Anna M. Borghi[1,2]*, *Chiara Fini*[1] and *Claudia Mazzuca*[1]

[1] Laboratory BALLAB (Body Action Language LAB), Department of Dynamic and Clinical Psychology and Health Studies, Sapienza University of Rome, Rome, Italy, [2] Institute of Cognitive Sciences and Technologies, Italian National Research Council, Rome, Italy

Keywords: abstract concepts, belief, abstract words, Theory of Mind, social interaction, concepts, metacognition, categorization

INTRODUCTION

This paper focuses on abstract concepts like "fantasy," "self-determination," and belief. Proficiently mastering abstract concepts and the words expressing them is a complex ability at the core of human cognition. Notably, we distinguish between "abstraction," i.e., the process leading to the formation of categories and order them hierarchically, and the related but independent notion of "abstractness," a characteristic of abstract concepts (Borghi, 2022). Scholars have a growing consensus that abstract concepts evoke sensorimotor experiences, although they are more detached from them than concrete concepts. Specifically, feature production tasks showed that, with abstract concepts, participants produce sensorimotor and interoceptive properties (Harpaintner et al., 2018; Banks et al., 2021). Rating tasks showed that participants judged that sensorimotor features and effectors such as the mouth and hand are involved in abstract concepts (Ghio et al., 2013; Villani et al., 2019). Finally, brain imaging evidence reveals that sensorimotor brain areas are recruited during abstract concept processing (Sakreida et al., 2013; Kiefer and Harpaintner, 2020). Typically, abstract concepts do not refer to single objects or entities but rather to situations, events, and feelings. Compared to concrete concepts, they evoke more interoceptive, inner bodily experiences and are more variable across individuals, languages, and cultures. Notably, concrete and abstract concepts are not dichotomously opposed; we can distinguish concrete and abstract categories into subkinds. Within concrete categories, the most widely investigated include artifacts, natural objects (Humphreys and Forde, 2001), and food (Rumiati and Foroni, 2016). As to abstract concepts, various sub-kinds exist too, like emotional (Ponari et al., 2018), numerical (Fischer and Shaki, 2018), social (Mellem et al., 2016), and Theory of Mind concepts (Desai et al., 2018; review: Conca et al., 2021). For each of these conceptual kinds, different dimensions—sensorimotor, interoceptive, linguistic, social—might be varyingly relevant in a flexible and contextual-dependent way (Borghi, 2022). For example, emotional concepts rely more on interoceptive experience (Connell et al., 2018; Villani et al., 2021b) and institutional concepts on linguistic and social ones, but expertise modulates the relevance of these dimensions (Villani et al., 2021a).

ABSTRACT CONCEPTS AND BELIEF

Belief During Abstract Concepts Acquisition and Use

Why might research on abstract concepts be relevant to literature on beliefs? We contend that it is significant for various reasons. First, because of the mechanism underlying abstract concepts' processing, which entails forming relevant beliefs on the conceptual content and the status of our own knowledge. Literature on categorization reveals that children and adults form relevant beliefs on categories, fostered by language generics (e.g., "sharks attack swimmers": Gelman et al., 2010; Gelman and Roberts, 2017). We propose that such beliefs are particularly relevant for forming and using abstract concepts. In our view, being more

OPEN ACCESS

Edited by:
Rüdiger J. Seitz, Heinrich Heine University of Düsseldorf, Germany

Reviewed by:
Rory Allen, Goldsmiths University of London, United Kingdom

***Correspondence:**
Anna M. Borghi
anna.borghi@uniroma1.it

Specialty section:
This article was submitted to Theoretical and Philosophical Psychology, a section of the journal Frontiers in Psychology

Received: *13 April 2022*
Accepted: *31 May 2022*
Published: *29 June 2022*

Citation:
Borghi AM, Fini C and Mazzuca C (2022) Abstract Concepts, Social Interaction, and Beliefs. Front. Psychol. 13:919808. doi: 10.3389/fpsyg.2022.919808

complex than concrete concepts, abstract ones involve a monitoring process (Borghi et al., 2019; Borghi, 2022); individuals assess the reliability of their knowledge. If we define beliefs as meaningful probabilistic representation (Seitz et al., 2018), then during conceptual processing, we would form a belief, which might not be explicit, about the conceptual content and the level and status of our knowledge of a given domain. While the outcome of this monitoring process is typically successful with concrete concepts, this is often not the case with abstract ones. We continuously seek evidence to confirm (and eventually confute) our beliefs, the character of which is dynamic and changeable. Because abstract concepts are complex and indeterminate in meaning, the beliefs concerning their content might be easily disconfirmed and require updating. Consistently, we typically feel more uncertain of our knowledge and understanding of abstract than concrete concepts. As recent evidence suggests, this feeling of uncertainty and low confidence can sometimes become explicit. A rating study demonstrated that people are less confident with abstract than concrete concepts (Mazzuca et al., 2022). Curiously, participants were not only less confident in their own knowledge, but they were also less trustful about experts' knowledge of the same domains.

We posit that a second mechanism showcasing the relationship between abstract concepts and beliefs enters into play when the monitoring process fails. We have called social metacognition (Borghi et al., 2018) the process occurring when we fail to find responses in ourselves and resort to others (Shea, 2018). Consistent evidence shows that participants performing an image-word guessing task tend to ask more for advice from others with abstract compared to concrete words; when performing a subsequent joint action task they show more inter-individual synchrony (Fini et al., 2021a). Furthermore, much evidence showed higher activation of the mouth motor system during abstract than concrete concepts acquisition and processing (review: Mazzuca et al., 2021; see also Ghio et al., 2013; Dreyer and Pulvermüller, 2018; Fini et al., 2021b; Reggin et al., 2021). One possible explanation for this activation is that we might prepare to ask others for information. When uncertain about word meaning, we might appeal to others.

We propose that this process might occur explicitly because we think that others might contribute to our knowledge, or it might be simply an automatic tendency. We might resort to others for various reasons. First, we might need them as experts who explain a concept or help us grasp it more in-depth. Forming beliefs about the authority and expertise of various people might help us select our possible interlocutor. Second, we might appeal to others because we struggle to understand their understanding of a concept or beliefs about its conceptual content. To illustrate, we might attempt to understand our interlocutor's notion of "freedom" to maximize the conversation's quality and mutual understanding. Notably, in some cases, people might violate these maxims and deliberately misinterpret others, especially when their ideological positions differ from their interlocutor's. Finally, we might rely on others because abstract concepts, having a less determined meaning, are more contestable than concrete concepts (Mazzuca and Santarelli, 2022). So, we might feel the need to negotiate their meaning with others. For example, imagine two scientists defining the notion of "belief." In this case, the two scientists will compare their beliefs on the notion to reach a consensus. Importantly, people might adopt different strategies to evaluate a source as reliable. For example, many people might consider a source reliable if the source agrees with them and unreliable in the opposite case. However, developmental literature on testimony reveals that, even if children tend to favor informants of their ingroup, they revoke their trust in familiar teachers if aware that they receive inaccurate information and select informants who do not flaunt confidence (Kominsky et al., 2016). Reliance on others is clearly more complex when the novel information pertains to issues that are the object of different ideologies.

Similarly, recent findings show that people are more uncertain about abstract concepts and tend to revolve more toward others. For instance, in a recent study, participants had to imagine a dialogue with an acknowledgment; they were prompted by a sentence including different kinds of abstract and concrete concepts (e.g., "I thought about destiny" vs. "I saw a lion") and had to write a possible response. With abstract concepts, and particularly with the more abstract among them, i.e., philosophical-religious concepts, participants asked more "why" and "how" questions—as opposed to "what" and "where" questions. In addition, they were keener to relaunch the conversation instead of closing it (e.g., "Explain me better") and used more words signaling uncertainty (e.g., "mmm") (Villani et al., 2022). This might be because the indeterminate character of the former might leave more space for the conversation to grow and for the exposition and possible alignment of idiosyncratic beliefs to emerge. Finally, Fini et al. (in preparation) found that during a virtual conversation, the psychological closeness between the interlocutors increased as a function of the importance attributed to the other's contribution to the dialogue. The other's contribution was especially relevant when the conversational topic was abstract. Conversing about an abstract topic might foster the creation of common beliefs; these processes of intellectual agreement might boost the overlapping between self-other representations.

To sum up, the study of abstract concepts might be significant for research on belief because during the use of abstract more than concrete concepts, we form a belief in the reliability and solidity of our knowledge, and because we seek to understand more the beliefs of others in given domains (see **Figure 1**—An example of the mechanisms described: the abstract concept "Self-determination").

The abstract concept of belief: There is another, probably more straightforward way studies on abstract concepts are relevant to the literature on beliefs. The notion of belief is inherently abstract, and those that are commonly called conceptual beliefs (Seitz and Angel, 2020) represent a sub-kind of abstract concepts. Research on abstract concepts can contribute to understanding how laypeople conceptualize the notion of belief. Notably, however, what people think about beliefs and how they actually use them are not necessarily strictly interrelated. In a study targeting 425 Italian abstract concepts, Villani et al. (2019) asked participants to evaluate them according to various dimensions, including imageability,

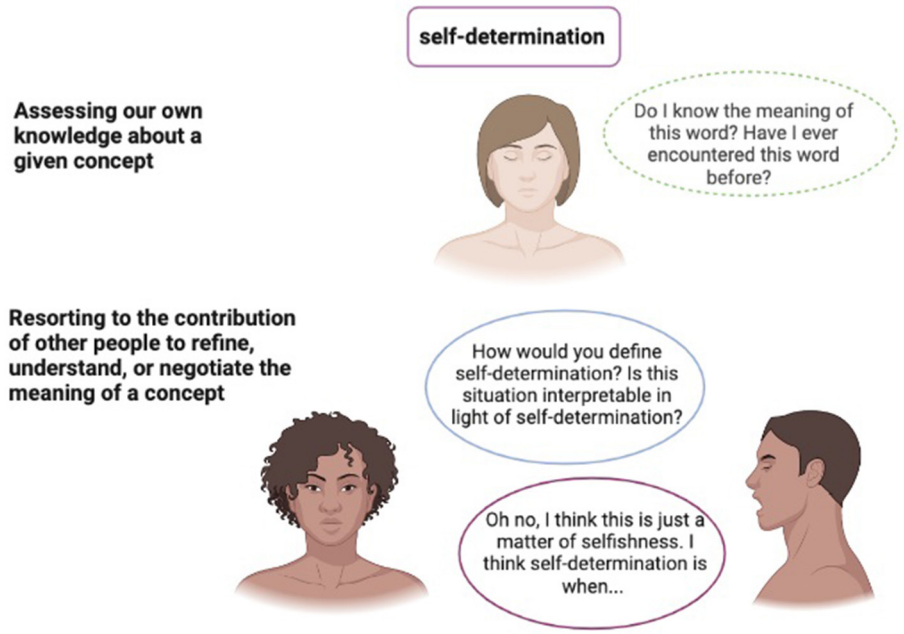


FIGURE 1 | The process of social metacognition.

contextual availability, activation of the five senses, interoception, metacognition, and sociality. A Principal Component Analysis revealed these concepts were characterized by three main dimensions (i.e., sensorimotor aspects, inner grounding, and abstractness∼concreteness), and four distinct clusters were identified. Among these, one cluster is predominantly composed of the more abstract among abstract concepts, which are more detached from sensorimotor and inner bodily experiences. We called these Philosophical-Spiritual abstract concepts. Notably, this cluster encompasses several mental states concepts, including "belief," together with concepts like "religion," and "faith." Further information on how belief is conceptualized comes from research on the neural underpinnings of concepts; for example, evidence shows that the brain representation of Theory of Mind concepts, like "belief," partially overlaps with the regions engaged by moral and emotional concepts (Desai et al., 2018).

CONCLUSION

To conclude, in this paper, we identified two main reasons why investigating abstract concepts can be useful for research on beliefs. First, we suggested that the mechanisms underlying abstract concepts processing and use involve multiple forms of beliefs. This is evident in the assessment of the solidity of our knowledge and the knowledge of our interlocutors. Second, we proposed that studies on abstract concepts can provide insights into how we represent the notion of belief in our brain and mind. We hope that future research will produce exciting new findings in this novel area.

AUTHOR CONTRIBUTIONS

AB drafted the paper. CM and CF revised and integrated it. All the authors conceived the article. All authors contributed to the article and approved the submitted version.

FUNDING

This paper is funded by Dr. Rüdiger Seitz, via the Volkswagen Foundation, Siemens Healthineers, and the Betz Foundation. All authors are funded by the H2020 project TRAINCREASE from social interaction to abstract concepts and words, grant proposal n. 952324. AB was funded also by the Sapienza Excellent Grant-Concepts in interaction with others and with ourselves: Abstractness in social interaction, metacognition and mind wandering, proposal n. RG12117A5D1EB0B3.

ACKNOWLEDGMENTS

We would like to thank all the members of BALLAB for the insightful discussions on this topic: Marta Arcovito, Federico Da Rold, Ilenia Falcinelli, Arthur-Henry Michalland, and Luca Tummolini.

REFERENCES

Banks, B., Wingfield, C., and Connell, L. (2021). Linguistic distributional knowledge and sensorimotor grounding both contribute to semantic category production. *Cogn. Sci.* 45, e13055. doi: 10.1111/cogs.13055

Borghi, A. M., Barca, L., Binkofski, F., Castelfranchi, C., Pezzulo, G., and Tummolini, L. (2019). Words as social tools: language, sociality and inner grounding in abstract concepts. *Phys. Life Rev.* 29, 120–153. doi: 10.1016/j.plrev.2018.12.001

Borghi, A. M., Barca, L., Binkofski, F., and Tummolini, L. (2018). Abstract concepts, language and sociality: from acquisition to inner speech. *Philos. Trans. R. Soc. B Biol. Sci.* 373, 20170134. doi: 10.1098/rstb.2017.0134

Borghi, A. M. (2022). Concepts for which we need others more: the case of abstract concepts. *Curr. direct. Psych. Sci.* doi: 10.1177/09637214221079625

Conca, F., Borsa, V. M., Cappa, S. F., and Catricalà, E. (2021). The multidimensionality of abstract concepts: a systematic review. *Neurosci. Biobehav. Rev.* 127, 474–491. doi: 10.1016/j.neubiorev.2021.05.004

Connell, L., Lynott, D., and Banks, B. (2018). Interoception: the forgotten modality in perceptual grounding of abstract and concrete concepts. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 373, 20170143. doi: 10.1098/rstb.2017.0143

Desai, R. H., Reilly, M., and van Dam, W. (2018). The multifaceted abstract brain. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 373, 20170122. doi: 10.1098/rstb.2017.0122

Dreyer, F. R., and Pulvermüller, F. (2018). Abstract semantics in the motor system? - An event-related fMRI study on passive reading of semantic word categories carrying abstract emotional and mental meaning. *Cortex* 100, 52–70. doi: 10.1016/j.cortex.2017.10.021

Fini, C., Era, V., Da Rold, F., Candidi, M., and Borghi, A. M. (2021a). Abstract concepts in interaction: The need of others when guessing abstract concepts smooths dyadic motor interactions. *R. Soc. Open Sci.* 8, 201205. doi: 10.1098/rsos.201205

Fini, C., Zannino, G. D., Orsoni, M., Carlesimo, G. A., Benassi, M., and Borghi, A. M. (2021b). EXPRESS: Articulatory suppression delays processing of abstract words: the role of inner speech. *Q. J. Exp. Psychol.* 75, 1343–54. doi: 10.1177/17470218211053623

Fischer, M. H., and Shaki, S. (2018). Number concepts: abstract and embodied. *Philos. Trans. R. Soc. Lond. B, Biol. Sci.* 373, 20170125. doi: 10.1098/rstb.2017.0125

Gelman, S. A., and Roberts, S. O. (2017). How language shapes the cultural inheritance of categories. *Proc. Natl. Acad. Sci.* 114, 7900–7907. doi: 10.1073/pnas.1621073114

Gelman, S. A., Ware, E. A., and Kleinberg, F. (2010). Effects of generic language on category content and structure. *Cogn. Psychol.* 61, 273–301. doi: 10.1016/j.cogpsych.2010.06.001

Ghio, M., Vaghi, M. M. S., and Tettamanti, M. (2013). Fine-grained semantic categorization across the abstract and concrete domains. *PLoS ONE* 8, e67090. doi: 10.1371/journal.pone.0067090

Harpaintner, M., Trumpp, N. M., and Kiefer, M. (2018). The semantic content of abstract concepts: A property listing study of 296 abstract words. *Front. Psychol.* 9:1748. doi: 10.3389/fpsyg.2018.01748

Humphreys, G. W., and Forde, E. M. (2001). Hierarchies, similarity, and interactivity in object recognition: "category-specific" neuropsychological deficits. *Behav. Brain Sci.* 24, 453–476; discussion 476–509. doi: 10.1017/S0140525X01004150

Kiefer, M., and Harpaintner, M. (2020). Varieties of abstract concepts and their grounding in perception or action. *Open Psychol.* 2, 119–137. doi: 10.1515/psych-2020-0104

Kominsky, J. F., Langthorne, P., and Keil, F. C. (2016). The better part of not knowing: virtuous ignorance. *Dev. Psychol.* 52, 31. doi: 10.1037/dev0000065

Mazzuca, C., Falcinelli, I., Michalland, A. H., Tummolini, L., and Borghi, A. M. (2022). Bodily, emotional, and public sphere at the time of COVID-19. An investigation on concrete and abstract concepts. *Psychol. Res.* 1–12. doi: 10.1007/s00426-021-01633-z

Mazzuca, C., Fini, C., Michalland, A. H., Falcinelli, I., Da Rold, F., Tummolini, L., et al. (2021). From Affordances to Abstract Words: The Flexibility of Sensorimotor Grounding. *Brain Sci.* 11, 1304. doi: 10.3390/brainsci11101304

Mazzuca, C., and Santarelli, M. (2022). Making it abstract, making it contestable: politicization at the intersection of political and cognitive science. *Rev. Philos. Psychol.* 1–22. doi: 10.1007/s13164-022-00640-2

Mellem, M. S., Jasmin, K. M., Peng, C., and Martin, A. (2016). Sentence processing in anterior superior temporal cortex shows a social-emotional bias. *Neuropsychologia* 89, 217–224. doi: 10.1016/j.neuropsychologia.2016.06.019

Ponari, M., Norbury, C. F., and Vigliocco, G. (2018). Acquisition of abstract concepts is influenced by emotional valence. *Dev. Sci.* 21. doi: 10.1111/desc.12549

Reggin, L. D., Muraki, E. J., and Pexman, P. M. (2021). Development of abstract word knowledge. *Front. Psychol.* 12, 2115. doi: 10.3389/fpsyg.2021.686478

Rumiati, R. I., and Foroni, F. (2016). We are what we eat: how food is represented in our mind/brain. *Psychon. Bull. Rev.* 23, 1043–1054. doi: 10.3758/s13423-015-0908-2

Sakreida, K., Scorolli, C., Menz, M. M., Heim, S., Borghi, A. M., and Binkofski, F. (2013). Are abstract action words embodied? An fMRI investigation at the interface between language and motor cognition. *Front. Hum. Neurosci.* 125. doi: 10.3389/fnhum.2013.00125

Seitz, R. J., and Angel, H. F. (2020). Belief formation–A driving force for brain evolution. *Brain Cogn.* 140, 105548. doi: 10.1016/j.bandc.2020.105548

Seitz, R. J., Paloutzian, R. F., and Angel, H.-F. (2018). From believing to belief: a general theoretical model. *J. Cogn. Neurosci.* 30, 1254–1264. doi: 10.1162/jocn_a_01292

Shea, N. (2018). Metacognition and abstract concepts. *Philos. Trans. R. Soc. Lond. B, Biol. Sci.* 373, 20170133. doi: 10.1098/rstb.2017.0133

Villani, C., D'Ascenzo, S., Borghi, A. M., Roversi, C., Benassi, M., and Lugli, L. (2021a). Is justice grounded? How expertise shapes conceptual representation of institutional concepts. *Psychol. Res.* 1–17. doi: 10.1007/s00426-021-01492-8

Villani, C., Lugli, L., Liuzza, M. T., and Borghi, A. M. (2019). Varieties of abstract concepts and their multiple dimensions. *Lang. Cogn.* 11, 403–430. doi: 10.1017/langcog.2019.23

Villani, C., Lugli, L., Liuzza, M. T., Nicoletti, R., and Borghi, A. M. (2021b). Sensorimotor and interoceptive dimensions in concrete and abstract concepts. *J. Mem. Lang.* 116, 104173. doi: 10.1016/j.jml.2020.104173

Villani, C., Orsoni, M., Lugli, L., Benassi, M., and Borghi, A. M. (2022). *Abstract and Concrete Concepts in Conversation. [Preprint]*. doi: 10.31219/osf.io/kv6dp

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Borghi, Fini and Mazzuca. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Behavioral Neuroscience

TYPE Hypothesis and Theory
PUBLISHED 06 October 2022
DOI 10.3389/fnbeh.2022.894219

Check for updates

OPEN ACCESS

EDITED BY
Igor Riecansky,
Slovak Academy of Sciences (SAS), Slovakia

REVIEWED BY
Michal Hajduk,
Comenius University, Slovakia
Philippe Bertrand,
Université Paris Cité, France

*CORRESPONDENCE
Rüdiger J. Seitz
seitz@neurologie.uni-duesseldorf.de

SPECIALTY SECTION
This article was submitted to Emotion Regulation and Processing, a section of the journal Frontiers in Behavioral Neuroscience

RECEIVED 11 March 2022
ACCEPTED 12 September 2022
PUBLISHED 06 October 2022

CITATION
Seitz RJ, Angel H-F, Paloutzian RF and Taves A (2022) Believing and social interactions: effects on bodily expressions and personal narratives. Front. Behav. Neurosci. 16:894219. doi: 10.3389/fnbeh.2022.894219

COPYRIGHT
© 2022 Seitz, Angel, Paloutzian and Taves. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Believing and social interactions: effects on bodily expressions and personal narratives

Rüdiger J. Seitz[1]*, Hans-Ferdinand Angel[2], Raymond F. Paloutzian[3] and Ann Taves[4]

[1]Department of Neurology, Centre of Neurology and Neuropsychiatry, LVR-Klinikum Düsseldorf, Medical Faculty, Heinrich-Heine-University Düsseldorf, Düsseldorf, Germany, [2]Institute of Catechetic and Pedagogic of Religion, Karl Franzens University Graz, Graz, Austria, [3]Department of Psychology, Westmont College, Santa Barbara, CA, United States, [4]Department of Religious Studies, University of California, Santa Barbara, CA, United States

The processes of believing integrate external perceptual information from the environment with internal emotional states and prior experience to generate probabilistic neural representations of events, i.e., beliefs. As these neural representations manifest mostly below the level of a person's conscious awareness, they may inadvertently affect the spontaneous person's bodily expressions and prospective behavior. By yet to be understood mechanisms people can become aware of these representations and reflect upon them. Typically, people can communicate the content of their beliefs as personal statements and can summarize the narratives of others to themselves or to other people. Here, we describe that social interactions may benefit from the consistency between a person's bodily expressions and verbal statements because the person appears authentic and ultimately trustworthy. The transmission of narratives can thus lay the groundwork for social cooperation within and between groups and, ultimately, between communities and nations. Conversely, a discrepancy between bodily expressions and narratives may cause distrust in the addressee(s) and eventually may destroy social bonds.

KEYWORDS

belief, trust, narratives, rituals, preferences, valuation, credition

Introduction

Believing conveys personal meanings that are constructed by means of perceptual and evaluative processes (Coltheart et al., 2011; Seitz et al., 2017). Believing processes also include predictive coding, which influences peoples' behavior as they make decisions. Whereas in philosophy beliefs are thought to be consciously held propositions (Stanford Encyclopedia), the concept of *credition* posits that beliefs are based on believing processes that are mostly nonconscious but that may become conscious when a person is believing (Angel et al., 2017). Accordingly, at the neuropsychological level believing can be

considered as a higher-order, integrative brain function similar to cognition and emotion (Angel et al., 2017; Angel, 2021). Notably, behavioral studies have revealed that the formation and updating of beliefs occur at a pre-linguistic level in non-human primates (Maravita and Iriki, 2004). However, humans can become aware of their beliefs and express their content and strength verbally (Oakley and Halligan, 2017; Seitz and Angel, 2020). Consequently, beliefs can be expected to play an important role in both verbal and non-verbal social interactions.

The findings and argument of this article lead to a novel perspective of the role of believing and beliefs in the shared realities of cultural dynamics that is underrepresented in recent literature (Kashima et al., 2018). They also raise the intriguing question of how communicating personal statements touches upon the as-yet not well-understood role of conscious awareness of belief contents in transmitting them from one person to other people. In order to explore these and related issues, let us begin with a look at the relationship between information processing and the formation and articulation of beliefs, trusting them, the human capacity to be conscious, and other aspects of human engagement that are rooted in credition–processes of believing. We will then be prepared to conceptualize the role of perceptual information processing, emotional valuation, and their appraisal in believing and decision making. Our discussion then shifts to examining the impact of believing on the generation of bodily expressions and verbal statements—which may be either intended or involuntary but are nevertheless interrelated in social communication. We conclude by describing these aspects of believing processes in relation to the dynamic evolution of social collaborations in ethnic groups, which may also apply to cultures and worldviews.

Belief formation, trust, and awareness

Information processing

Living beings process a great deal of information about physical objects in their environment. Importantly, at the neurophysiological level, they process this information very rapidly. This speed of transmission is part of what enabled them to evolve. In the same manner, they also rapidly process information about events, which are things perceived by an observer as a change in the environment with a beginning and an end (Zacks and Tversky, 2001; Asprem and Taves, 2021). The information about objects and events has to be weighed as beneficial or aversive, and must allow a person to react both fast and appropriately. When positive emotions are involved, affirmative beliefs become manifest; this is in contrast to when negative emotions are involved, which render objects and events as threatening, irritating, or disgusting (Seitz et al., 2018). Such processes, which involve the complex interaction of the perception of objects and assessing their positive or negative value and emotional tone, are intimate to meaning-making and remaking (Paloutzian and Mukai, 2017). They constitute the fundamental ground of the processes of believing at the neuropsychological level, and they extrapolate to the psychological, social, and cultural levels as well, with increasing complexity at each step. Therefore, across levels of analysis, a belief is a meaning that has been made and stored in memory (Seitz et al., 2022).

In addition to the pre-linguistic type of belief formation and updating as noted above, humans also process verbal information. From birth onwards, verbal information is provided by caregivers, and later with increasing age by many other people. Also, verbal information is often presented in a ritual fashion through nursery rhymes, songs, fairy tales, and stories. Such narratives are spoken or written accounts of events that are connected and loaded with positive emotions. Often, such narratives can function as the basis for the intuitive generation of conceptual beliefs about a personal self, a family, a social group, and a community, as well as place, time, morals, justice, and many other aspects of social life (Belzen, 2010a,b; Zaidel, 2019). From an evolutionary perspective, it is interesting that ritual activities and play behavior have a number of features in common and are widespread in non-human animals (Mori, 2020).

Repetition and trusting

People typically believe that what they have perceived is accurate and true; they intuitively trust their perceptions, because they are processed easily and concerning the environment typically are true (Brashier and Marsh, 2020). However, if the events are below 200 ms, and, therefore, cannot be stored in memory correctly, claims about the perception are typically not accurate (Bear et al., 2017). Thus, there is a close relationship between believing and trust. Trust has been defined in different fields of study—personality theory, sociology, economics, social psychology—and summarized as an individual's belief and willingness to act (Lewicki and Tomlinson, 2014). As such, trust comprises a number of social-cognitive dimensions such as competence, integrity, predictability, compassion, compatibility, etc. (Kappmeier, 2016). For example, it has recently been shown that repetitive stimulation induces people to trust their perceptions, which can lead to an illusory truth-effect (Fazio and Sherry, 2020). These findings suggest that, although someone may believe one or another element of environmental information, it may require

a number of converging observations before a person trusts a situation or another person.

Awareness

The processes of believing occur so rapidly that information perceived from the environment is integrated with internal emotional loadings prior to conscious awareness (Wegner, 2003; Seitz et al., 2009; Park and Tallon-Baudry, 2014). The speed of this integration is similar to that of the generation of a simple motor action; for example, as when the flexing of an index finger is initiated below conscious awareness (Libet, 1985; Hallett, 2016). In a similar way, when developing expectations and preferences humans typically rely on relatively stable conceptual beliefs without being aware of them (Williams, 2020). This finding supports the notion that understanding how unconscious knowledge works is fundamental to understanding human thought processes and mentation more generally (Augusto, 2010).

Even so, the content of thoughts and beliefs may enter conscious awareness and allow an individual to give a verbal account of what he or she believes (Oakley and Halligan, 2017). The neural processes underlying belief formation and updating have been shown to demand a phylogenetic expansion of brain functions that enable people to make verbal statements that begin with "I believe . . ." (Seitz and Angel, 2020). The ability to express what one believes has been hypothesized to be the prerequisite for auto-reflexive as well as interpersonal belief evaluation (Langdon and Coltheart, 2000; Seitz, 2022). However, most behavior is not pre-thought or generated by "reason". But as soon as someone becomes aware of an intended action, the person is capable of voluntarily modulating the behavior up to a certain point, as has been shown experimentally (Filipović et al., 2000). This capability is reflected in the common German expression "sich beherrschen" (keep calm). It means that a person who might spontaneously act with high internal drive in a possibly exaggerated manner has a limited time window in which to calm down and voluntarily suppress aversive acts, so that the behavior turns out to be appropriate for the circumstance. The need for humans and other animals to modulate their actions so that they are consistent with the norms and values of the individual's social network requires that a valuative process be part of the processes of believing.

Valuation of information

Probabilistic

When humans interact with objects or other people, they intuitively develop an affective attitude that reflects the putative beneficial or aversive impact of the encounter (Seitz et al., 2009; Prochnow et al., 2013). The resulting probabilistic perceptive-emotional accounts are the basis for the person's predictions of future events as well asfor context-related adaption of his or her behavior (**Figure 1**). Accordingly, the emotional valence renders the perceived object or event personally relevant, and shapes what a person intuitively uses for behavioral control. The neural representation that provides this tight neural link between the information that has been perceived and the prediction that determines the selection of a subsequent action has a probabilistic character and, thus, may be considered a consequence of believing processes, i.e., a belief (Seitz et al., 2018). Similarly, narratives about how an individual comes to belong to or be part of a certain group—such as a family, ethnic tribe, or regionally defined group—exert a strong influence when they are presented in ritual acts. And because of their strong affective components, rituals stabilize social behavior within and across generations—a phenomenon that has also been described in non-human animals.

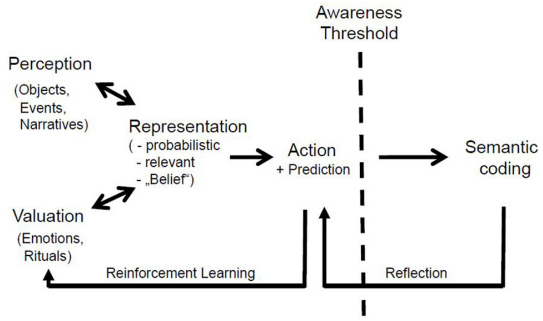


FIGURE 1
The neural processes underlying believing below and above a person's conscious awareness. The processes on the left evolve fast, within the range of milliseconds in cortico-subcortical brain structures, allowing for the formation and updating of beliefs and corresponding action generation. Perception pertains to environmental information, whereas valuation mediates the emotional valence. These processes can be the object of empirical neuroscience research. The neural processes on the right occur in the realm of conscious awareness and capitalize on verbalized information which can become the object of a person's reflection or appraisal. Note that the participation in rituals provides an immediate emotional loading to such stereotypic events that are instructed by corresponding narratives. These putative processes have stimulated epistemic theories in the humanities since antiquity.

Speed and affect

In addition to its probabilistic nature, emotionally-laden information may be differentially detected as, e.g., when its speed is in the range of milliseconds (below human conscious awareness). For example, people could not detect fearful-looking faces when they were flashed for 33 ms, but they did detect the faces when presented for 67 ms (Pessoa et al., 2006). Also, in

a backward masking task, the occipitotemporal N170 electrical potential was enhanced when people observed facial expressions that were categorized as emotional, which suggests that their brains were processing information from the faces without conscious awareness (Smith, 2012). In another study of how long it takes for someone to detect a fearful face when presented subliminally and supraliminally, it was found that detecting a fearful face occurred at approximately 260–300 ms after presentation (Pegna et al., 2008). This finding corresponds to the notion that perceptual awareness emerges at about 200 ms, with modality-specific negative changes in the brain at 120–200 ms and a later modality-independent positive potential at about 300 ms (Koivisto and Revonsuo, 2010; Marti and Dehaene, 2017; Dembski et al., 2021). The amygdala seems to be of critical importance for this to occur since it is said to coordinate the activity of cortical networks during the early evaluation of the biological significance of affective visual stimuli (Pessoa and Adolphs, 2010). The activity of the amygdala has also been found to be modulated in people who had to look at photos with stimuli placed to the left or right of pictures that depicted emotionally loaded fearful human faces (Straube et al., 2010; Kim et al., 2016). These data suggest that under certain conditions, subconscious processing of inferences may overcome the threshold of awareness, e.g., when there are temporally extended or repetitive observations. In fact, there is evidence that low-level inferences may occur fast and below conscious awareness, whereas high level conscious inferences integrate information across different sensory modalities and larger spatial scales and, therefore, take more time (Olcese et al., 2018).

Emotions shape what a person remembers because emotional cues play a fundamental role in gating relevant information and suppressing non-relevant information. For example, positively valenced stimuli improve prospective memory performance such that events with a strong emotional loading can be retrieved from memory more vividly than neutral events (Hostler et al., 2018; Kensinger and Ford, 2020). In addition, in empirical studies with more than 20,000 individuals, it was found that attitudes based on emotions were relatively fixed and decayed less over time if the emotions were positive (Rocklage and Luttrell, 2021). Thus, the emotional loading of the inherently ambiguous information about objects and events enhances the probability of its encoding as a personal imagination in memory.

Fluidity

Nevertheless, owing to their probabilistic nature, beliefs are fluid and can be updated upon new evidence depending on the subjective weighing of the previous or new information (Angel and Seitz, 2017; Seitz et al., 2018; Kube and Rozenkrantz, 2021). Even positive perceptions can turn into negative perceptions. An example is the change of attitude towards wetness of the skin. Thermoregulatory behavior is known to depend both on peripheral sensors that communicate their information to the brain, as well as on temperature sensing within the brain (Tan and Knight, 2018). Specifically, individuals were found to perceive warm-wet and neutral-wet stimuli as significantly less wet than cold-wet stimuli on their skin, although the stimuli had the same moisture content (Filingeri et al., 2014). Likewise, on a hot summer day, very wet skin due to a lot of perspiration can cause someone to feel uncomfortable and possibly some disgust, whereas similarly wet skin as the result of a cool bath may be perceived as joyful and refreshing. Yet, on the evening of such a summer day, a bath of similar temperature may be experienced as unpleasant and to be avoided, very much similar to a bath on a cold and overcast day. These examples are consistent with the observation that attitudes, preferences, and values are not absolute. Rather, their coding of valence seems to follow a relative scale (Vlaev et al., 2011; Pischedda et al., 2020).

Evolving of valuation

Early age

Before children begin to speak and learn words for the objects and events around them, they learn to interact nonverbally with other people. They learn to recognize emotional facial expressions and communicative gestures. Thus, children learn to make sense of communicative acts and nonverbal gestures from first-hand observation (Harris et al., 2018). In addition, they imitate the motor acts they observe and learn that they get praise for doing this well (Piaget, 1978). In doing so, they learn to associate their own facial expressions with their emotional feelings. Evidence for this is illustrated by a field experiment in which it was found that young people who reported more intense experiences of fear and happiness were more accurate in recognizing facial expressions of fear and happiness by the early age of 5 years (Buchanan et al., 2010). Children have also been reported to understand the content of other minds through social and communicative interactions with others, which requires that they compare their own perspective to that of others (Tomasello, 2018). As children learn to acquire such information, which comes from multimodal external sources, they apparently reason about how trustworthy the information they are receiving is (Harris et al., 2018). Children thereby develop a sense of trust in their representations of their environment, of which two important aspects are a sense of authorship and causal inference.

Inferences and conceptual beliefs

It has been shown that humans track the likelihood that their inferences are correct such that probabilistic learning and

estimating confidence in what has been learned are intimately related (Meyniel et al., 2015). Although confidence increases with the number of observations, children have been found to be prone to set aside their own prior convictions and defer to informants for social reasons when they are presented with unexpected or counterintuitive but still credible testimony (Harris et al., 2018). Thus, beliefs can be modified in view of new information that is valued higher than previous information. Specifically, social reasoning appears to be valenced higher than one's own sense of trustworthiness (Harris et al., 2018). It may seem remarkable that personal appreciation of a social relationship is intuitively valued so strongly as to override one's individual stance. But people in close relationships are likely to be connected by similar beliefs and values, which allows them to maintain common meaning systems (Andersen and Przybylinski, 2018). Such commonality seems to involve predictions about the other person's most likely behavior, including the non-verbal mentalizing capacity called "theory of mind" (Bird and Viding, 2014). For example, in the cortical areas that have been associated with the "theory of mind", personally familiar faces have been shown to evoke stronger responses than faces of famous people who happen to be known but not personally (Gobbini et al., 2004). Further, people show an inherent tendency toward intuitive prosociality, as social learning involves areas ascribed to the so-called social brain such as the ventral medial prefrontal cortex, in addition to areas involved in self-relevant learning (Lengersdorff et al., 2020).

Children probably acquire conceptual beliefs in an intuitive fashion from early on. For example, nursery tales, narratives about ritual acts, and the proper prayers in religious families are communicated to children regularly and shape their beliefs and worldviews. Consistent with this, it has been argued that a sense of morality could emerge in a developmental system in which children's early capacities are shaped by interpersonal engagement (Carpendale and Hammond, 2016). Only later, upon explicit reasoning about such conceptual beliefs, will the information in these communications be brought into conscious awareness so that the person can begin to talk about their beliefs and what their implications mean to them (**Figure 2**).

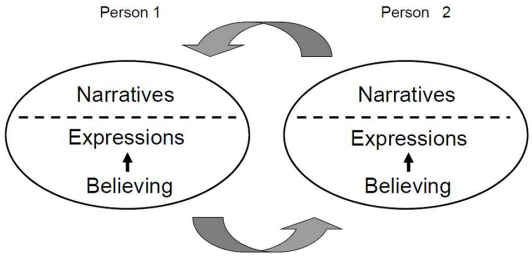


FIGURE 2
Interpersonal exchange of personally relevant information. Believing can result in bodily, non-verbal expressions and/or verbal statements executed spontaneously below the realm of conscious awareness (dotted line). Narratives include spontaneous personal statements and refined conceptual beliefs. Note that interpersonal interactions involving non-verbal and/or verbal expressions and narratives also occur within groups of people and among members of communities.

Communicating the contents of beliefs

Mentalizing and self-narratives

Beliefs about objects and events are behaviorally highly important and inadvertently affect bodily movements, as can be observed in facial expressions, gestures, and other coordinated or uncoordinated actions (Dael et al., 2012). It typically seems almost impossible to suppress these spontaneous bodily reactions because their purpose is to immediately convey behaviorally relevant information to others (**Figure 2**). An example of such a bodily reaction is facial mimicry. Facial mimicry occurs when someone observes the facial expression of an emotion and has a strong impulse to express the same emotion (Müller et al., 2019). Thus, emotionally loaded attitudes and beliefs can facilitate bodily expressions of feelings of which the individual may be unaware. However, people can also become aware of their beliefs and express their content semantically—a process that has been referred to as internal broadcasting (Oakley and Halligan, 2017). Such intrapersonal communication occurs in several modes including inner dialogue and self-talk (Oles et al., 2020). These inner dialogues are characterized by so-called "I positions", which represent familiar elements of first-person experience (Langland-Hassan, 2021). Such first-person semantic expressions are "I find", which is primarily emotional, "I think", which sounds rational but is nevertheless vague, and "I believe", which conveys a clear, unambiguous personal stance (Seitz and Angel, 2020). Expressions such as "I like . . .", "I want . . ." and "I believe . . ." may also convey an affirmative attitude. In contrast, verbal expressions that convey aversive information or a negative attitude are "I fear ...", "I hate . . .", "I am angry about. . .", and "I am disgusted . . .". Both the affirmative and the negative meanings of these and similar expressions are stored in long-term memory and can be retrieved at later times. In both cases, the individual entertains either an inclination towards or an aversion to the perceptual-emotional accounts. These internal propositions enable a person to perform abstract thought and executive functions, and thereby support meta-cognition.

Narrating our personal past connects us to ourselves, our families, our communities, and our cultures (Fivush et al., 2011). Conceptual beliefs as expressed in personal narratives are comprised of information about autobiographical memory that underpins constructs of personal self and agency that were created in non-conscious systems (Oakley and Halligan, 2017).

Thus, the so-called narrative or autobiographical self represents a self-image that consists of memories of the past and intentions about the future, constituted from the various stories that people have been told (Gallagher, 2000). Based on this, self-talk supports self-reinforcement, self-management, self-criticism, and social assessment (Oles et al., 2020; Paloutzian et al., 2021a). In other words, humans are in the position to reflect upon the contents of their beliefs (**Figure 1**). Evaluation of such internal narratives help someone assess the degree to which they reflect reality and are trustworthy, and to modify them with respect to relevant environmental conditions. An implication is that people can reflect on their behavior and act for reasons they can specify.

Honest or not?

Communicating to others *via* narratives may be intended to convey personally relevant information in order to enhance interpersonal relationships, possibly for the sake of socio-ecology allowing for common goals or actions (Romano et al., 2021). We tend to perceive someone as authentic when the person's verbal reports and spontaneous bodily expressions including the facial expressions are congruent with each other (Franz et al., 2021). Then, we tend to trust what the person says. However, someone with a manipulative or deceptive intent can covertly modify the narrative as detailed in the truth-default theory (Levine, 2022). In these cases, the speaker's spontaneous bodily expressions may convey an intent that differs from his or her statements or narratives. If the recipients of the communication detect such a discrepancy, they may perceive the speaker's message as false and the person as unreliable. In particular, a delay and discrepancy between the speaker's thought-based verbal expressions and spontaneous bodily expressions may signal that the speaker is anxious, unsure, thoughtless, or deceitful. Concerning narratives, there is an intriguing question about whether emotional valence is communicated by the vocal tuning of verbal expressions, or is conveyed by words themselves. On this issue, a correlation analysis of the assessment of more than 1,400 English words found that abstract words receive higher ratings for affective associations, including valence and arousal, than concrete words (Vigliocco et al., 2014). This finding supports the notion that abstract words are more emotionally valenced than concrete words (Montefinese, 2019). Likewise, language metaphors probably covertly influence people's reasoning even when different options of how to phrase something are available (Thibodeau and Boroditsky, 2013).

Although narratives evolve rapidly, they can extend over variable lengths of time, depending on the complexity of their content. Neural processing evolves fast enough to allow for the information to be decoded. For example, in audiovisual recognition of vowels coordinated oscillations in speech areas including the inferior frontal gyrus can be detected (Lange et al., 2013). The neural activity changes in relation to the complexity of semantic tasks. For example, it was found recently that neural oscillations encoded endogenously generated linguistic content that surpassed exogenous stimulus-driven timing and rhythm information (Kaufeld et al., 2020). This finding is consistent with the notion that these bioelectric changes could reflect computations related to how humans and other animals infer structure and meaning from acoustic signals (Kaufeld et al., 2020). In other words, the data suggest that the human brain is capable of creating a meaning from a sequence of acoustic stimuli that goes beyond a single stimulus. This may be the basis for abstraction, which allows us to make "bigger" meanings out of the initially specific meanings, i.e., in conceptual psychological terms, to make more global meanings out of lesser situational meanings (Park, 2010). This is also of relevance to the notion of transcendence in language processing (Mesulam, 1990).

Individuals in groups

Identity and bonding

Groups allow their members to behave differently (Barrett et al., 2001). For example, primates of the same species do not necessarily act in an identical way to all members of their species. But they do demonstrate in-group bias, favoritism, and altruism, as well as out-group prejudice, disfavor, and lack of help of an "other". This pattern of behavior is identical to what Tajfel described in elaborating the social identity theory (Outten et al., 2018). In classic research, when subjects (children, adults, teenagers, adults) originating from different cultures are randomly divided into groups (and know that they are assigned to their group by random chance), they still show an in-group bias and out-group prejudice (Tajfel, 1981). The in-group ("Us") is better, smarter, prettier, and the out-group ("Them") is stupid, worse, and more ugly. Such findings suggest that primate sociality has evolved so that it is based on bonded social relationships (Dunbar and Shultz, 2007). Bondedness is an explicitly emotional experience that integrates group perception with an internal affective state or intuitive emotional valence. This is related to what has been called relational beliefs (Seitz and Angel, 2020), i.e., a person who feels emotionally connected to another person or persons spontaneously tends to trust them.

An integrated trust model unifies the existing literature on the multidimensionality of trust, and allows us to explore the role of trust in social collaboration as well as the bases of intergroup conflict or tension, as illustrated in ethnic discrimination (Kappmeier et al., 2019). It also is consistent with Tajfel's social identity theory and can account for the reactions of majority group members towards minorities in different societies (Outten et al., 2018). There is also a close relationship between social identity and the impact of collective memory such that their

combined residue can foster either intra-group trust or inter-group conflict (Kappmeier and Mercy, 2019). Thus, someone in Group A claims to have "owned" or "created" this desired object but believes that the "other" in Group B deserves only that undesired thing. Likewise, social comparison processes that result in attitudes and behaviors of the "I am better than you" sort are manifest not only in competition for staples like food and sex, but also in competition for the sake of social recognition and the superiority of oneself within one's own group. These processes occur without necessarily being clear or explicit to an individual.

In accordance with these intuitive processes, narratives have been stated to support social and cultural structures (Oakley and Halligan, 2017). Narratives can convey the contents of similar conceptual beliefs such as family, honesty, fraternity, equality, charity, etc. to different people. Because narratives are stored in the memory of individual subjects and can be retrieved from their memory at later time points (Seitz et al., 2022), they are fundamental for the conservation of conceptual beliefs in social groups and societies. Narratives provide the reason for and occasion to engage in ritual acts that are practiced in families, social groups, and communities (Schnell, 2012; Gelfand et al., 2020; Mori, 2020). People thereby develop their social identity narratives of ethnic culture as well as individual self-concept (Knight et al., 2018). Concurrent with this, children are taught and learn how to behave in their social environment.

Cultures and values

Religions are, amongst other things, cultures (Cohen, 2009). One observation consistent with the above argument is that there is an association between the profession of religious devotion and greater trusting behavior (Norenzayan and Shariff, 2008). This association may occur as beliefs in a morally concerned god may stabilize prosocial norms within the culture even in the absence of social monitoring mechanisms. Such stabilizing may occur at the neural level, in that religious beliefs were found to activate regions within a network related to mentalizing of intent and emotion, abstract semantics, and imagery (Kapogiannis et al., 2009). In related research, the comparison of religious and non-religious subjects did not reveal any differences in these activations—in accordance with the notion that religiosity is integrated into cognitive processes and brain networks used in social cognition (Boyer, 2003). Extending the above notions, there are potent models to explain how ethnic views expand among groups and extrapolate to explain the acquisition of similar views and subsequent related behavior in other cultures (Galesic and Stein, 2019). Even so, it is possible for someone's cultural orientation to change over time as a function of their experiences with and membership in multiple groups, in addition to their normal age-related developmental changes (Knight et al., 2018).

In any case, language is considered to be an inadequate medium for describing inner emotional experiences and communicating them (Dunbar and Shultz, 2007). The reasons are twofold, namely that the speaker needs to become aware of his/her emotional experience and needs to know how to express this experience clearly in words. These descriptions of personal experience also need to have some meaning for the listener. Such communication of meaning is probably straight-forward for basic emotions like happiness, sadness, anger, and fear. However, more complex feelings or "higher" emotional values like empathy, forgiveness, and altruism may not be easily understood or straight-forwardly shared spontaneously. Instead, they may need to be explained by more elaborate verbal descriptions or perhaps be accompanied by a positive emotional descriptor. In this context the concept of shared reality is important. There is lot of evidence that communicators fine-tune their statements in an effort to align them with the attitudes of those to whom they are speaking. Doing this in turn shapes their recall (Higgins et al., 2021) and has been said to promote interpersonal closeness and epistemic certainty.

Extending the above argument further, certain moral values may be considered as higher order emotions and may function in a way similar to them. For example, the feeling of empathy is highly value-laden and implies accepting another person and his or her difficulties in a manner similar to accepting oneself. A common illustration is the moral values codified in religions, such as the Ten Commandments, which provide a guideline for how to behave properly. These and similar teachings are transmitted among people across generations and reflect not only stable language use but are also suited to guide certain behavior according to their norms. Thus, people can reflect on their thoughts, wishes, and actual actions in light of these norms, and thereby become responsible for their actions. At the neural level, a study in which the participants viewed scenes evocative of moral emotions showed that the orbital and rostral medial prefrontal cortex and the cortex along the superior temporal sulcus are involved in mediating the above noted value-related events (Moll et al., 2002). Processes such as trusting, forgiving, and believing matter because humans make attributions about these properties and respond accordingly (Paloutzian et al., 2021b). Problems arise between parties when there is an inconsistency between what one says and what one does when verbal behavior and overt actions are discrepant (ibid). Collaboration can only re-start in small, reciprocal, trust-inducing steps (ibid). This means that the actual experience with the counterpart matches what he or she believes about the counterpart being of particular relevance for international and cross-cultural issues (Schoorman et al., 2007).

Discussion

The neural processes that afford belief formation, believing, and the updating of beliefs occur spontaneously in the time domain of milliseconds. As summarized in **Figure 1**, belief formation includes the integration of information coming from the environment and attribution of emotional value, with both aspects resulting in personal probabilistic representations. This model accounts also for the formation of socially adaptive beliefs that are sensitive to social rewards and punishments (Williams, 2020). Accordingly, beliefs are intimately coupled with subjective experience prior to complex processing of prediction of a behavioral outcome and to awareness of the incoming information as suggested recently (Key et al., 2022). Belief updating occurs by means of reinforcement learning *via* cortico-subcortical circuits when actual and predicted information match, whereas new information of high subjective relevance is able to induce a change in the belief (**Figure 1**). Therefore, the cerebral networks that are involved allow for the storage of beliefs in memory (Seitz et al., 2022). This is consistent with the notion that cerebral representations are memories that are localized in neural networks and, when activated, enable access to this stored information (Wood and Grafman, 2003).

In a very similar hypothesis, experienced events were labeled as event knowledge (Taves and Asprem, 2016). Even though event recognition and other processes are occurring, most brain processes are not accompanied by any discernable changes in subjective awareness (Halligan and Oakley, 2021). But people can retrieve stored information from memory, whereby it then enters their conscious awareness (**Figure 1**). This retrieval is a critical prerequisite for a person to be able to semantically phrase what he or she is believing. This can typically be done by implicit or deliberate self-talk or a prayer. Either way, people can reflect on their beliefs and sharpen their awareness of the information. This reflection probably corresponds to the notion of the belief evaluation systems (Coltheart et al., 2011; Sugiura et al., 2015), which may be explained by invoking the concept of event models. Event models are constructed from the point of view of the person who perceives the entities and functional relations involved in understanding a specific state of affairs (Radvansky and Zacks, 2017). As a result, they contain information that the person considers relevant regarding spatiotemporally located entities (agents and objects) and establish the structural and linking relations between them as he or she understands them in light of their previous experience [i.e., in light of plausible types of events (event schemas) and their own particular memories of past events (other event models)]. Also, relations that link objects and events, which include the causes and consequences of events, play a crucial role in the way the model is structured, linked to other events, and retrieved on later occasions. In fact, upon reflection people can modify their behavior so that it deviates from the predictions based on beliefs only. Beyond that, belief evaluation enables humans to communicate what they believe to other people (Oakley and Halligan, 2017). Consequently, exploring the neural principles of belief formation and updating is central to the research discipline of social cognitive neuroscience (Lieberman, 2010).

In the concept of credition, believing is a fundamental brain function that links emotional valence to sensory perceptions, rendering them personally relevant and memorable (Angel et al., 2017; Seitz et al., 2018, 2022; Seitz and Angel, 2020; Angel, 2021). In fact, emotion signals have been shown to enhance processing efficiency and competitive strength of emotionally significant events through gain control mechanisms mediated in the amygdala and interconnected prefrontal cortical areas (Pourtois et al., 2013). By this means, emotions become fundamental to the self-regulation of behavior (Peil, 2014), although they may change over one's lifetime. For example, toys, food, and drinks that infants and children love can be undesirable to adults. Conversely, the personal relevance of objects or events can be modified by diseases. For example, patients handicapped by a disabling disease of the body may still have a positive perspective on life in a way that may seem impossible for a healthy person. Thus, valence may inadvertently be changed by external events, which can result in an update or even dismissal of a hitherto held belief (Angel and Seitz, 2017). Moreover, brain diseases leading to neuropsychological deficits and psychopathological disorders have been shown to result in the formation of abnormal beliefs that can cause inadequate or even aversive behavior which can undermine social bonds (Connors and Coltheart, 2011; Seitz, 2022).

Probably because of the emotional and rapidly evolving nature of underlying neural processes, the processes of believing take place below a person's awareness and, thus, outside his/her reach. This becomes obvious in social interactions in which a person judges his/her counterpart and *vice versa*. Humans are known to rapidly develop an intuition or belief about whether to trust another person and how to react to him/her (Potthoff and Seitz, 2015). Such primal beliefs influence our spontaneous bodily expressions, as has been found in facial mimicry and bodily movements (**Figure 2**). Both are expressions of non-verbal communication (Dael et al., 2012). However, beliefs may enter conscious awareness—probably in a graded fashion rather than in an all or none manner for the different sensory modalities. Their content then can be phrased verbally, rehearsed internally, and communicated as narratives to others (**Figure 2**). It is a specific human capability that narratives underlying conceptual thinking and believing cannot only be transmitted *via* speech but can also be written down and transferred to other people as scripts, letters, or books (Belzen, 2010a,b). Such documents can be read, reflected on, and re-read, allowing for new associations and novel creative thoughts. Written concepts also support social

memory. Thereby, narratives turn out to be fundamental for the autobiographical self and the formation of social groups.

As to the bases for making predictions from such narratives, humans are in a position to explore whether their actions concur with norms, rules, and expectations of other people or whether they offend them. Having these options corresponds to what has been called to act based upon reasons (Proust, 2003). In so doing, people become "responsible" for their actions. This does not exclude that they may flexibly manipulate group-mates' behavior to tactically deceive them, as has been shown in experimental food competitions in primates (Hall and Brosnan, 2017). Humans may also intentionally deceive people, such as when there is a discrepancy between their pre-thought verbal statements and their spontaneous motor expressions. In this connection, there are neurophysiological and neuroanatomical bases for cognitive and affective theory of mind, with interpersonal and intrapersonal dimensions that humans can use to determine when cheaters need to be punished (Westby, 2014). In fact, humans are highly capable of detecting whether someone's verbal and non-verbal communication are consistent or inconsistent with each other. For human behavior, these different possibilities are accounted for by the cultural brain hypothesis. This hypothesis posits that brains have been selected by evolution for their ability to store and manage information that was acquired through social and asocial learning (Muthukrishna et al., 2018). Consistent with this idea, many components of language, including extra-linguistic meaning systems and the communication of symbolic meaning, have neurobiological roots that go back millions of years in evolutionary time (Zaidel, 2019).

The data are consistent with the notion that our capacity to use language creatively enables us to gain awareness of the mental worlds of other people, and that we can communicate our own imaginative play, creative narratives, original thoughts, arguments, and feelings to them (Markl, 2002). In essence, the multi-level analysis presented in this article appears capable of bridging the gaps between the level of neural systems to the behavioral level in individuals to the social level. We assume that probabilistic processes at the neural level and increased probability in a stepwise fashion as we go up to the behavioral and social levels. Nevertheless, it is important to remember that the relations between symbols and content can be quite variable across different cultures. For example, the association of *white* with *joy* and *black* with *grief* is a Western tradition, with associations in the opposite directions in Asia. Also in Western communities, shaking one's head means "no" and nodding means "yes", but these head movements convey opposite meanings in other cultures. Similarly, there are complex patterns of language evolution with respect to different ethnicities that involve adopting, keeping, and replacing vocabularies and grammars (Das et al., 2016). Relatedly, ratings of the degree of affect in neutral faces have been shown to not be neutral; they are instead loaded with different levels of ambiguity, and thus may yield important differential psychological consequences (Schneider et al., 2016). These findings raise interesting issues (yet to be solved) about the concurrence, discrepancy, and ambiguity of our verbal and non-verbal communication. In any case, human intelligence appears to be a combination and enhancement of properties found in non-human primates including mentalizing (theory of mind), imitation, and learning from verbal testimony (Roth and Dicke, 2005; Harris et al., 2018).

There is much psychological evidence that supports the proposition that majority views are held with stronger confidence and expressed more quickly than are minority views, regardless of any social pressure to conform (Koriat et al., 2016). Thus, social consensus plays a causal role in supporting and enhancing a person's confidence in beliefs, opinions, and attitudes (ibid). Further, social influence is involved when one attempts to either gain social approval or avoid social isolation. But when we consider real-world groups and the issues between them (men vs. women, blacks vs. whites, Middle-Eastern Muslims vs. Western Jews, and Christians, ... the list is endless), with actual fighting and lethal confrontations, we can understand not only why there is intergroup conflict but also how the tendency humans have towards outgroups has its roots in our genetic makeup from eons of evolution (Paloutzian et al., 2021b). Even so, if we humans can evolve inclinations to trust, including trusting our enemies (in graded mutual and reciprocal steps, so that it is possible for the process to work), we may evolve out of group conflict as "built in" to our genes towards contact and collaboration with all humans as one group, so that everybody can love everybody instead of being afraid of them. Ultimately, if everybody would just sit down and talk about their processes of believing, we would learn that we are more ike each other instead of the various ways that we differ.

Quite unexpectedly, it was found that affective content is highly relevant in abstract thoughts and conceptual beliefs (Montefinese, 2019). For example, religious beliefs have been shown to be maintained by prayer and ritual acts but not by deductive or inductive reasoning (Atran and Norenzayan, 2004; Feierman, 2009). Today, many people have greater confidence in their scientific beliefs than in their religious beliefs, although similar patterns of justification have been described for both kinds of believing (Harris and Corriveau, 2020). However, the comfort and support provided by religious organizations may grow when people experience more harshness, when coping resources begin to diminish, and when environmental pressures demand a greater effort (Seryczynska et al., 2021). Because the adults' perception of the relation between religion and science is heavily shaped by their sociocultural contexts, the relation between religiosity and the valuation of science varies profoundly between different countries (Payir et al., 2021). This does not preclude that the contents of different beliefs, such as political or religious

beliefs, may be reported to be equivalent but not identical (Oviedo and Szocik, 2020).

In conclusion, the notions of belief and believing are complex cognitive constructs similar to culture and consciousness that may be amenable to naturalistic exploration in an evolutionary framework (Singer, 2019).

Data availability statement

The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Author contributions

RS: designing, drafting, and editing. H-FA, RP, and AT: drafting and editing. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by Frontiers Science Publisher.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Andersen, S. M., and Przybylinski, E. (2018). Shared reality in interpersonal relationships. *Curr. Opin. Psychol.* 23, 42–46. doi: 10.1016/j.copsyc.2017.11.007

Angel, H.-F. (2021). *Credition: Processes and Functions of Believing (website)*. Austria: University of Graz.

Angel, H.-F., Oviedo, L., Paloutzian, R. F., Runehov, A. L. C., and Seitz, R. J. (2017). *Processes of Believing: The Acquisition, Maintenance and Change in Creditions*. Heidelberg: Springer.

Angel, H. F., and Seitz, R. J. (2017). Violation of expectation as matter for the believing process. *Front. Psychol.* 8:772. doi: 10.3389/fpsyg.2017.00772

Asprem, E., and Taves, A. (2021). "Event model analysis," in *The Routledge Handbook of Research Methods in the Study of Religion*, 2nd edition, eds S. Engler, and M. Stausberg (New York: Taylor and Francis, Routledge).

Atran, S., and Norenzayan, A. (2004). Religion's evolutionary landscape: counterintuition, commitment, compassion, communion. *Behav. Brain Sci.* 27, 713–730. doi: 10.1017/s0140525x04000172

Augusto, L. M. (2010). Unconscious knowledge: a survey. *Adv. Cogn. Psychol.* 6, 116–141. doi: 10.2478/v10053-008-0081-5

Barrett, J. L., Richert, R. A., and Driesenga, A. (2001). God's beliefs versus mother's: the development of nonhuman agent concepts. *Child Dev.* 72, 50–65. doi: 10.1111/1467-8624.00265

Bear, A., Fortgang, R. G., Bronstein, M. V., and Cannon, T. D. (2017). Mistiming of thought and perception predicts delusionality. *Proc. Natl. Acad. Sci. U S A* 114, 10791–10796. doi: 10.1073/pnas.1711383114

Belzen, J. A. (2010a). *Towards Cultural Psychology of Religion: Principles, Approachs, Applications*. Dordrecht: Springer.

Belzen, J. A. (2010b). Religion and self: notions from a cultural psychological perspective. *Pastoral Psychol.* 59, 199–409. doi: 10.1007/s11089-009-0204-z

Bird, G., and Viding, E. (2014). The self to other model of empathy: providing a new framework for understanding empathy impairments in psychopathy, autism and alexithymia. *Neurosci. Biobehav. Rev.* 47, 520–532. doi: 10.1016/j.neubiorev.2014.09.021

Boyer, P. (2003). Religious thought and behaviour as by-products of brain function. *Trends Cogn. Sci.* 7, 119–124. doi: 10.1016/s1364-6613(03)00031-7

Brashier, N. M., and Marsh, E. J. (2020). Judging truth. *Ann. Rev. Psychol.* 71, 499–515. doi: 10.1146/annurev-psych-010419-050807

Buchanan, T. W., Bibas, D., and Adolphs, R. (2010). Associations between feeling and judging the emotions of happiness and fear: findings from a large-scale field experiment. *PLoS One* 5:e10640. doi: 10.1371/journal.pone.0010640

Carpendale, J. I., and Hammond, S. I. (2016). The development of moral sense and moral thinking. *Curr. Opin. Pediatr.* 28, 743–747. doi: 10.1097/MOP.0000000000000412

Cohen, A. B. (2009). Many forms of culture. *Am. Psychol.* 64, 194–204. doi: 10.1037/a0015308

Coltheart, M., Langdon, R., and McKay, R. (2011). Delusional belief. *Annu. Rev. Psychol.* 62, 271–298. doi: 10.1146/annurev.psych.121208.131622

Connors, M. H., and Coltheart, M. (2011). On the behaviour of senile dementia patients vis-à-vis the mirror: ajuriagurra, strejilevitch and tissot (1963). *Neuropsychologia* 49, 1679–1692. doi: 10.1016/j.neuropsychologia.2011.02.041

Dael, N., Mortillaro, M., and Scherer, K. R. (2012). Emotion expression in body action and posture. *Emotion* 12, 1085–1101. doi: 10.1037/a0025737

Das, R., Wexler, P., Pirooznia, M., and Elhaik, E. (2016). Localizing ashkenazic jews to primeval villages in the ancient iranian lands of ashkenaz. *Genome Biol. Evol.* 8, 1132–1149. doi: 10.1093/gbe/evw046

Dembski, C., Koch, C., and Pitts, M. (2021). Perceptual awareness negativity: a physiological correlate of sensory consciousness. *Trends Cogn. Sci.* 25, 660–670. doi: 10.1016/j.tics.2021.05.009

Dunbar, R. I. M., and Shultz, S. (2007). Evolution in the social brain. *Science* 317, 1344–1347. doi: 10.1126/science.1145463

Fazio, L. K., and Sherry, C. L. (2020). The effect of repetition on truth judgments across development. *Psychol. Sci.* 1, 1–11. doi: 10.1177/0956797620939534

Feierman, J. R. (2009). "How some major component of religion could have evolved by natural selection?," in *The Biological Evolution of Religious Mind and Behavior. The Frontiers Collection*, eds E. Voland, and W. Schiefenhövel (Heidelberg: Springer), 51–66

Filingeri, D., Fournet, D., Hodder, S., and Haenith, G. (2014). Why wet feels wet? A neurophysiological model of human cutaneous wetness sensitivity. *J. Neurophysiol.* 112, 1457–1469. doi: 10.1152/jn.00120.2014

Filipović, S. R., Jahanshahi, M., and Rothwell, J. C. (2000). Cortical potentials related to the nogo decision. *Exp. Brain Res.* 132, 411–415. doi: 10.1007/s002210000349

Fivush, R., Habermas, T., Waters, T. E. A., and Zaman, W. (2011). The making of autobiographical memory: intersections of culture, narratives and identity. *Int. J. Psychol.* 46, 321–345. doi: 10.1080/00207594.2011.596541

Franz, M., Nordmann, M. A., Rehagel, C., Schäfer, R., Müller, T., and Lundqvist, D. (2021). It is in your face - alexithymia impairs facial mimicry. *Emotion* 21, 1537–1549. doi: 10.1037/emo0001002

Galesic, M., and Stein, D. L. (2019). Statistical physics models of belief dynamics: theory and empirical tests. *Physica A* 519, 275–294. doi: 10.1016/j.physa.2018.12.011

Gallagher, I. (2000). Philosophical conceptions of the self: implications for cognitive science. *Trends Cogn. Sci.* 4, 14–21. doi: 10.1016/s1364-6613(99)01417-5

Gelfand, M. J., Caluori, N., Jackson, J. C., and Taylor, M. K. (2020). The cultural evolutionary trade-off of ritualistic synchrony. *Phil. Trans. R. Soc. B* 375:20190432. doi: 10.1098/rstb.2019.0432

Gobbini, M. I., Leibenluft, E., Santiago, N., and Haxby, J. V. (2004). Social and emotional attachment in the neural representation of faces. *Neuroimage* 22, 1628–1635. doi: 10.1016/j.neuroimage.2004.03.049

Hall, K., and Brosnan, S. F. (2017). Cooperation and deception in primates. *Infant. Behav. Dev.* 48, 38–44. doi: 10.1016/j.infbeh.2016.11.007

Hallett, M. (2016). Physiology of free will. *Ann. Neurol.* 80, 5–12. doi: 10.1002/ana.24657

Halligan, P. W., and Oakley, D. A. (2021). Giving up on consciousness as the ghost in the machine. *Front. Psychol.* 12:571460. doi: 10.3389/fpsyg.2021.571460

Harris, P. L., and Corriveau, K. H. (2020). Beliefs of children and adults in religious and scientific phenomena. *Curr. Opin. Psychol.* 40, 20–23. doi: 10.1016/j.copsyc.2020.08.003

Harris, P. L., Koenig, M. A., Corriveau, K. H., and Jaswal, V. K. (2018). Cognitive foundations of learning from testimony. *Annu. Rev. Psychol.* 69, 251–273. doi: 10.1146/annurev-psych-122216-011710

Higgins, E. T., Rossignac-Milon, M., and Echterhoff, G. (2021). Shared reality: from sharing-is-believing to merging minds. *Curr. Direct. Psychol. Sci.* 30, 103–110. doi: 10.1177/0963721421992027

Hostler, T. J., Wood, C., and Armitage, C. J. (2018). The influence of emotional cues on prospective memory: a systematic review with meta-analyses. *Cogn. Emot.* 32, 1578–1596. doi: 10.1080/02699931.2017.1423280

Kapogiannis, D., Barbey, A. K., Su, M., Zamboni, G., Krueger, F., Grafman, J., et al. (2009). Cognitive and neural foundations of religious belief. *Proc. Natl. Acad. Sci. U S A* 106, 4876–4881. doi: 10.1073/pnas.0811717106

Kappmeier, M. (2016). Trusting the enemy - towards a comprehensive understanding of trust in intergroup confluct. *Peace Conflict J. Peace Psychol.* 22, 134–144doi: 10.1037/pac0000159

Kappmeier, M., Guenoun, B., and Campbell, R. (2019). They are us? The mediating effects of compatibility-based trust on the relationship between discrimination and overall trust. *N. Z. J. Psychol.* 48, 97–105.

Kappmeier, M., and Mercy, A. (2019). The long road from cold war to warm peace: building shared collective memory through trust. *J. Soc. Political Psychol.* 7, 525–555. doi: 10.5964/jspp.v7i1.328

Kashima, Y., Bratanova, B., and Peters, K. (2018). Social transmission and shared realtiy in cultural dynamics. *Curr. Opin. Psychol.* 23, 15–19. doi: 10.1016/j.copsyc.2017.10.004

Kaufeld, G., Bosker, H. R., ten Oever, S., Alday, P. M., Meyer, A. S., Martin, A. E., et al. (2020). Linguistic structure and meaning organize neural oscillations into a content-specific hierarchy. *J. Neurosci.* 40, 9467–9475. doi: 10.1523/JNEUROSCI.0302-20.2020

Kensinger, E. A., and Ford, J. H. (2020). Retrieval of emotional events from memory. *Annu. Rev. Psychol.* 71, 251–272. doi: 10.1146/annurev-psych-010419-051123

Key, B., Zalucki, O., and Brown, D. J. (2022). A first principles approach to subjective experience. *Front. Syst. Neurosci.* 16:756224. doi: 10.3389/fnsys.2022.756224

Kim, M. J., Solomon, K. M., Neta, M., Davis, F. C., Oler, J. A., Mazzulla, E. C., et al. (2016). A face versus non-face context influences amygdala responses to masked fearful eye whites. *Soc. Cogn. Affect. Neurosci.* 11, 1933–1941. doi: 10.1093/scan/nsw110

Knight, G. P., Safa, M. D., and White, R. M. B. (2018). Advancing the assessment of cultural orientiation: a developmental and contextual framework of multiple psychological dimensions and social identities. *Dev. Psychopathol.* 30, 1867–1888. doi: 10.1017/S095457941800113X

Koivisto, M., and Revonsuo, A. (2010). Event-related brain potential correlates of visual awareness. *Neurosci. Biobehav. Rev.* 34, 922–934. doi: 10.1016/j.neubiorev.2009.12.002

Koriat, A., Adiv, S., and Schwarz, N. (2016). Vies that are shared with others are expressed with greater confidence and greater fluency independent of any social influence. *Personal. Soc. Psychol. Rev.* 20, 176–193. doi: 10.1177/1088868315585269

Kube, T., and Rozenkrantz, L. (2021). When beliefs face reality: an integrative review of belief updating in mental health and illness. *Perspect. Psychol. Sci.* 16, 247–274. doi: 10.1177/1745691620931496

Langdon, R., and Coltheart, M. (2000). The cognitive neuropsychology of delusions. *Mind Lang.* 15, 184–218. doi: 10.1111/1468-0017.00129

Lange, J., Christian, N., and Schnitzler, A. (2013). Audio-visual congruency alters power and coherence of oscillatory activity within and between cortical areas. *Neuroimage* 79, 111–120. doi: 10.1016/j.neuroimage.2013.04.064

Langland-Hassan, P. (2021). Inner speech. *Wiley Interdiscip. Rev. Cogn. Sci.* 12:e1544. doi: 10.1002/wcs.1544

Lengersdorff, L. I., Wagner, I. C., Lockwood, P. I., and Lamm, C. (2020). When implicit prosociality trumps selfishness: the neural valuation systems underpins more optimal choices when learning to avoid harm to others than to oneself. *J. Neurosci.* 40, 7286–7299. doi: 10.1523/JNEUROSCI.0842-20.2020

Levine, T. R. (2022). Truth-default theory and the psychology of lying and deception detection. *Curr. Opin. Psychol.* 47:101380. doi: 10.1016/j.copsyc.2022.101380

Lewicki, R. J., and Tomlinson, E. C. (2014). "Trust, trust development and trust repair," in *The Handbook of Conflict Resolution*, 3rd edition, P. T. Coleman, M. Deutsch, and E. C. Marcus (New York, USA: Jossey-Bass), 104–136

Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behav. Brain Sci.* 8, 529–566. doi: 10.1017/S0140525X00044903

Lieberman, M. D. (2010). "Social cognitive neuroscience," in *Handbook of Social Psychology*, 5th edition, S. T. Fiske, D. T. Gilbert, and G. Lindsey (New York: John Wiley Sons Inc.), 143–193.

Maravita, A., and Iriki, A. (2004). Tools for the body (schema). *Trends Cogn. Sci.* 8, 79–86. doi: 10.1016/j.tics.2003.12.008

Markl, H. (2002). Man's place in nature: evolutionary past and genomic future. *J. Mol. Biol.* 319, 869–876. doi: 10.1016/s0022-2836(02)00331-5

Marti, S., and Dehaene, S. (2017). Discrete and continuous mechanisms of temporal selection in rapid visual streams. *Nat. Commun.* 8:1955. doi: 10.1038/s41467-017-02079-x

Mesulam, M. M. (1990). Large-scale neurocognitive networks and distributed processing for attention, language and memory. *Ann. Neurol.* 28, 597–613. doi: 10.1002/ana.410280502

Meyniel, F., Schlunegger, D., and Dehaene, S. (2015). The Sense of confidence during probabilistic learning: a normative account. *PLoS Comput. Biol.* 11:e1004305. doi: 10.1371/journal.pcbi.1004305

Moll, J., de Oliveira-Souza, R., Eslinger, P. J., Bramati, I. E., Mourao-Miranda, J., Andreiuolo, P. A., et al. (2002). The neural correlates of moral sensitivity: a functional magnetic resonance imaging investigation of basic and moral emotions. *J. Neurosci.* 22, 2730–2736. doi: 10.1523/JNEUROSCI.22-07-02730.2002

Montefinese, M. (2019). Semantic representation of abstract and concrete words: a minireview of neural evidence. *J. Neurophysiol.* 121, 1585–1587. doi: 10.1152/jn.00065.2019

Mori, I. (2020). "Introducing ritual play and belief," in *Evolution and Early Human Societies*. Cambridge, UK: University of California at St. Barbara Libraries.

Müller, T., Schäfer, R., Hahn, S., and Franz, M. (2019). Adults' facial reaction to affective facial expressions of children and adults. *Int. J. Psychophysiol.* 139, 33–39. doi: 10.1016/j.ijpsycho.2019.01.001

Muthukrishna, M., Doebell, M., Chudek, M., and Henrich, J. (2018). The cultural brain hypothesis: how culture drives brain expansion, sociality and life history. *PLoS Comput. Biol.* 14:e1006504. doi: 10.1371/journal.pcbi.1006504

Norenzayan, A., and Shariff, A. F. (2008). The origin and evolution of religious prosociality. *Science* 322, 58–62. doi: 10.1126/science.1158757

Oakley, D. A., and Halligan, P. W. (2017). Chasing the rainbow: the non-conscious nature of being. *Front. Psychol.* 8:1924. doi: 10.3389/fpsyg.2017.01924

Olcese, U., Oude Lohuis, M. N., and Pennartz, C. M. A. (2018). Sensory processing across conscious and nonconscious brain states: from single neurons to distributed networks for inferential representation. *Front. Syst. Neurosci.* 12:49. doi: 10.3389/fnsys.2018.00049

Oles, P. K., Brinthaupt, T. M., Dier, R., and Polak, D. (2020). Types of inner dialogues and functions of self-talk: comparisons and implications. *Front. Psychol.* 11:227. doi: 10.3389/fpsyg.2020.00227

Outten, H. R., Lee, T., Costa-Lopes, R., Schmitt, M. T., and Vala, J. (2018). Majority group members' negative reactions to future demographic shifts depend on the perceived legitimacy of their status: findings from the united states and portual. *Front. Psychol.* 9:79. doi: 10.3389/fpsyg.2018.00079

Oviedo, L., and Szocik, K. (2020). Religious and other beliefs: how much specificity? *Sage Open* 1:21582440198. doi: 10.1177/2158244019898849

Paloutzian, R. F., and Mukai, K. J. (2017). "Believing, remembering and imagining: the roots and fruits of meanings Mmade and remade," in *Processes of Believing: The Acquisition, Maintenance and Change in Creditions*, eds H.-F. Angel, L. Oviedo, R. F. Paloutzian, A. L. C. Runehov, and R. J. Seitz (Dordrecht, Heidelberg: Springer), 39–49.

Paloutzian, R. F., Seitz, R. J., and Angel, H.-F. (2021a). "The processes of believing and communicating with the unseen," in *Human Interaction With the Divine, the Sacred and the Deceased: Psychological, Scientific and Theological Perspectives*, eds T. Plante, and G. Schwartz (London and New York: Taylor and Francis, Routledge).

Paloutzian, R. F., Sagir, Z., and Shults, F. L. (2021b). "Modelling reconciliation and peace processes," in *Lessons From Syrian War Refugees and World War II*, 13th edition, eds K. P. Clemens, and S. Y. Lee (New York: Routledge), 225–242

Park, C. L. (2010). Making sense of the meaning literature: an integrative review of meaning making and its effects on adjustment to stressful life events. *Psychol. Bull.* 136, 257–301. doi: 10.1037/a0018301

Park, H.-D., and Tallon-Baudry, C. (2014). The neural subjective frame: from bodily signals to perceptual consciousness. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 369:20130208. doi: 10.1098/rstb.2013.0208

Payir, A., Davoodi, T., Cui, K. Y., Clegg, J. M., Harris, P. L., Corriveau, K., et al. (2021). Are high levels of religiosity inconsistent with a high valuation of science? Evidence from the united states, china and iran. *Int. J. Psychol.* 56, 216–227. doi: 10.1002/ijop.12701

Pegna, A. J., Landis, T., and Khateb, A. (2008). Electrophysiological evidence for early non-conscious processing of fearful facial expressions. *Int. J. Psychophysiol.* 70, 127–136. doi: 10.1016/j.ijpsycho.2008.08.007

Peil, K. T. (2014). Emotion: the self-regulatory sense. *Glob. Adv. Health Med.* 3, 80–108. doi: 10.7453/gahmj.2013.058

Pessoa, L., and Adolphs, R. (2010). Emotion processing and the amygdala: from a "low road" to "many roads" of evaluating biological significance. *Nat. Rev. Neurosci.* 11, 773–783. doi: 10.1038/nrn2920

Pessoa, L., Japee, S., Sturman, D., and Ungerleider, L. G. (2006). Target visibility and visual awareness modulate amygdala responses to fearful faces. *Cereb. Cortex* 16, 366–375. doi: 10.1093/cercor/bhi115

Piaget, J. (1978). *Behaviour and Evolution*. New York: Routledge.

Pischedda, D., Palminteri, S., and Coricelli, G. (2020). The effect of counterfactual information on outcome value coding in medial prefrontal and cingulate cortex: from an absolute to a relative neural code. *J. Neurosci.* 40, 3268–3277. doi: 10.1523/JNEUROSCI.1712-19.2020

Potthoff, D., and Seitz, R. J. (2015). Role of the first and second person perspective for control of behaviour: understanding other people's facial expressions. *J. Physiol. Paris* 109, 191–200. doi: 10.1016/j.jphysparis.2015.12.003

Pourtois, G., Schettino, A., and Vuilleumier, P. (2013). Brain mechanisms for emotional influences on perception and attention: what is magic and what is not. *Biol. Psychol.* 92, 492–512. doi: 10.1016/j.biopsycho.2012.02.007

Prochnow, D., Kossack, H., Brunheim, S., Müller, K., Wittsack, H.-J., Markowitsch, H.-J., et al. (2013). Processing of subliminal facial expressions of emotion: a behavioral and fMRI study. *Soc. Neurosci.* 8, 448–461. doi: 10.1080/17470919.2013.812536

Proust, J. (2003). "How voluntary are minimal actions?," in *Voluntary Action - Brains - Minds and Sociality*, eds S. Maassen, W. , Prinz, and G. Roth (Oxford, UK: Oxford University press), 202–219.

Radvansky, G. A., and Zacks, J. M. (2017). Event boundaries in memory and cognition. *Curr. Opin. Behav. Sci.* 17, 133–140. doi: 10.1016/j.cobeha.2017.08.006

Rocklage, M. D., and Luttrell, A. (2021). Attitudes based on feelings: fixed or fleeting? *Psychol. Sci.* 32, 364–380. doi: 10.1177/0956797620965532

Romano, A., Giardini, F., Columbus, S., de Kwadsteniet, E. W., Kisfalusi, D., Triki, Z., et al. (2021). Reputation and socio-ecology in humans. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 376:20200295. doi: 10.1098/rstb.2020.0295

Roth, G., and Dicke, U. (2005). Evolution of the brain and intelligence. *Trends Cogn. Sci.* 9:52. doi: 10.1016/j.tics.2005.03.005

Schneider, I. K., Veenstra, L., van Harreveld, F., Schwarz, N., and Koole, S. I. (2016). Let's not be indifferent about neutrality: neutral ratings in the international affective picture system (IAPS) mask mixed affective responses. *Emotion* 16, 426–430. doi: 10.1037/emo0000164

Schnell, T. (2012). Spirituality with and without religion. *Arch. Psychol. Rel.* 34, 33–62. doi: 10.1163/157361212X644495

Schoorman, F. D., Mayer, R. C., and Davis, J. H. (2007). An integrative model of organizational trust: past, present and future. *Acad. Manage. Rev.* 32, 344–354. doi: 10.5465/amr.2007.24348410

Seitz, R. J. (2022). Beliefs - a challenge in neuropsychological disorders. *J. Neuropsychol.* 6, 21–37. doi: 10.1111/jnp.12249

Seitz, R. J., and Angel, H.-F. (2020). Belief formation - a driving force for brain evolution. *Brain Cogn.* 140:105548. doi: 10.1016/j.bandc.2020.105548

Seitz, R. J., Angel, H. F., and Paloutzian, R. F. (2022). Bridging the gap between believing and memory functions. *Eur. J. Psychol. Psych. Arch.* doi: 10.23668/psycharchives.5421

Seitz, R. J., Franz, M., and Azari, N. P. (2009). Value judgments and self-control of action: the role of the medial frontal cortex. *Brain Res. Rev.* 60, 368–378. doi: 10.1016/j.brainresrev.2009.02.003

Seitz, R. J., Paloutzian, R. F., and Angel, H. F. (2017). Processes of believing. Where do they come from? What are they good for? *F1000Res.* 5:2573. doi: 10.12688/f1000research.9773.2

Seitz, R. J., Paloutzian, R. F., and Angel, H. F. (2018). From believing to belief: a general biomathematical model. *J. Cogn. Neurosci.* 30, 1254–1264. doi: 10.1162/jocn_a_01292

Seryczynska, B., Oviedo, L., Roszak, P., Saarelainen, S.-M. K., Inkilä, H., Albaladejo, J. T., et al. (2021). Religious capital as a central factor in coping with the covid-19 clues from an international survey. *Eur. J. Sci. Theol.* 17, 67–81.

Singer, W. (2019). A naturalistic approach to the hard problem of consciousness. *Front. Syst. Neurosci.* 13:58. doi: 10.3389/fnsys.2019.00058

Smith, M. L. (2012). Rapid processing of emotional expressions without conscious awareness. *Cereb. Cortex* 22, 1748–1760. doi: 10.1093/cercor/bhr250

Straube, T., Dietrich, C., Mothes-Lasch, M., Mentzel, H.-J., and Miltner, W. H. R. (2010). The volatility of the amygdala response to masked fearful eyes. *Hum. Brain Mapp.* 31, 1601–1608. doi: 10.1002/hbm.20960

Sugiura, M., Seitz, R. J., and Angel, H.-F. (2015). Models and neural bases of the believing process. *J. Behav. Brain Sci.* 5, 12–23. doi: 10.4236/jbbs.2015.51002

Tajfel, H. (1981). *Human Groups and Social Categories: Studies in Social Psychology*. Cambridge: Cambridge University Press.

Tan, C. L., and Knight, Z. A. (2018). Regulation of body temperature by the nervous system. *Neuron* 98, 31–48. doi: 10.1016/j.neuron.2018.02.022

Taves, A., and Asprem, E. (2016). Experience as event: event cognition and the study of (religious) experiences. *Rel. Brain Behav.* 7, 1–44. doi: 10.1080/2153599X.2016.1150327

Thibodeau, P. H., and Boroditsky, L. (2013). Natural language metaphors covertly influence reasoning. *PLoS One* 8:e52961. doi: 10.1371/journal.pone.0052961

Tomasello, M. (2018). How children come to understand false beliefs: a shared intentionality account. *Proc. Natl. Acad. Sci. U S A* 115, 8491–8498. doi: 10.1073/pnas.1804761115

Vigliocco, G., Kousta, S. T., Della Rosa, P. A., Vinson, D. P., Tettamami, M., Devlin, J. T., et al. (2014). The neural representation of abstract words: the role of emotion. *Cereb. Cortex* 24, 1767–1777. doi: 10.1093/cercor/bht025

Vlaev, I., Chater, N., Steward, N., and Brown, G. D. A. (2011). Does the brain calculate value? *Trends Cogn. Sci.* 15, 546–554. doi: 10.1016/j.tics.2011.09.008

Wegner, D. M. (2003). The mind's best trick: how we experience conscious will. *Trends Cogn. Sci.* 7, 65–69. doi: 10.1016/s1364-6613(03)00002-0

Westby, C. S. (2014). Social neuroscience and theory of mind. *Folia Phoniatr. Logop.* 66, 7–17. doi: 10.1159/000362877

Williams, D. (2020). Socially adaptive belief. *Mind Lang.* 1, 1–22. doi: 10.1111/mila.12294

Wood, J. N., and Grafman, J. (2003). Human prefrontal cortex: processing and representational perspectives. *Nat. Rev.* 4, 139–147. doi: 10.1038/nrn1033

Zacks, J. M., and Tversky, B. (2001). Event structure in perception and cognition. *Psychol. Bull.* 127, 3–21. doi: 10.1037/0033-2909.127.1.3

Zaidel, D. W. (2019). Coevolution of language and symbolic meaning: co-opting meaning underlying the initial arts in early human culture. *WIREs Cogn. Sci.* 11:e1520. doi: 10.1002/wcs.1520

frontiers | Frontiers in Psychology

TYPE Opinion
PUBLISHED 17 November 2022
DOI 10.3389/fpsyg.2022.949131

Check for updates

OPEN ACCESS

EDITED BY
Rüdiger J. Seitz,
Heinrich Heine University of Düsseldorf, Germany

REVIEWED BY
Neil Van Leeuwen,
Georgia State University, United States

*CORRESPONDENCE
Michiel van Elk
m.van.elk@fsw.leidenuniv.nl

SPECIALTY SECTION
This article was submitted to Theoretical and Philosophical Psychology, a section of the journal Frontiers in Psychology

RECEIVED 20 May 2022
ACCEPTED 15 July 2022
PUBLISHED 17 November 2022

CITATION
van Elk M (2022) Proximate and ultimate causes of supernatural beliefs. *Front. Psychol.* 13:949131. doi: 10.3389/fpsyg.2022.949131

COPYRIGHT
© 2022 van Elk. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Proximate and ultimate causes of supernatural beliefs

Michiel van Elk*

Institute of Psychology, Leiden University, Leiden, Netherlands

KEYWORDS

supernatural beliefs, cultural learning, dualism, agency detection, cross-cultural, replication, cognitive science of religion (CSR), psychology of religion and spirituality

Introduction

In the 2000s with the discovery of the so-called *God-spot*—a brain region that was suggested to be involved in the experience of God (Biello, 2007)—the field of neurotheology came to flourish, according to which supernatural beliefs are engrained in our brain. At the same time, other researchers have pointed out the relevance of socio-cultural factors for the learning and proliferation of supernatural beliefs (Norenzayan and Gervais, 2013), in line with the view that ultimately religion evolved through a process of cultural evolution, thereby fostering in-group cohesion and cooperation (Norenzayan et al., 2016). Still others have argued that religion primarily fulfills an epistemic need to understand and predict the world (Kay et al., 2010) and that it provides a palliative mechanism to cope with the fear of death (Vail et al., 2010).

Which of these viewpoints is right? What are the proximate and ultimate mechanisms that help us to understand why some people believe in supernatural phenomena, like an afterlife, spirit communication or a soul, whereas others don't? In this perspective paper I will provide a critical examination of the existing literature on this topic, especially in light of the so-called replication crisis: many published findings in the scientific literature turned out not to be replicable (Nosek et al., 2015). This was mainly related to questionable research practices, underpowered studies, lack of independent replication studies and the file-drawer problem and similar concerns have haunted the psychology and cognitive science of religion as well (van Elk et al., 2015; Charles et al., 2019). Therefore, in the Religious Replication Project (Hoogeveen and van Elk, 2018), over the past years we set out to assess the replicability of key findings in the field, by conducting direct replication studies of existing findings, registered report studies and large-scale cross-cultural replication studies. In this review I will specifically focus on what we learned about the (1) proximate cognitive mechanisms underlying supernatural beliefs, (2) the psychological functions subserved by supernatural beliefs, and (3) socio-cultural mechanisms contributing to the proliferation of supernatural beliefs (see Figure 1). I will end by discussing the implications of these different mechanisms for our understanding of the nature of supernatural beliefs and how they come about.

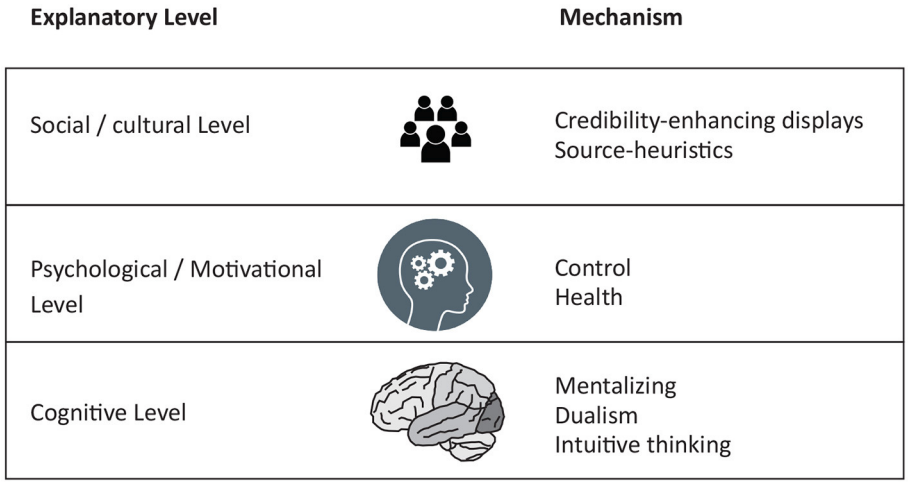


FIGURE 1
Overview of the different explanatory levels **(Left)** and the hypothesized mechanisms **(Right)** that have been proposed to account for the emergence and proliferation of supernatural beliefs.

Cognitive mechanisms underlying supernatural beliefs

Different proximate cognitive mechanisms have been proposed in the literature to underlie supernatural beliefs, including the ability to mentalize (Barrett, 2000), dualistic reasoning (Bering et al., 2005) and intuitive thinking (Boyer, 2001).

Mentalizing

The ability to apply theory of mind reasoning has been suggested to be a necessary prerequisite for enabling belief in an anthropomorphic supernatural agent and it has been found for instance that personal prayer to God is associated with the activation of brain regions involved in mentalizing (Schjoedt et al., 2009). Hyper-mentalizing, i.e., the tendency to attribute intentions to natural phenomena such as thunderstorms and earthquakes, has also been associated with an increased tendency to believe in supernatural and paranormal phenomena (Willard and Norenzayan, 2013). By using correlational designs, across several studies we were also able to show that stronger supernatural beliefs were associated with a stronger bias for illusory agency detection by using perceptual decision making tasks in which participants were required to indicate whether a human agent was visible in a display or not (van Elk, 2013, 2015). It has also been found that mentalizing deficits, e.g., as observed in people scoring high on the autism spectrum, are negatively related to belief in a personal god (Norenzayan et al., 2012). We replicated this finding in a large-scale cross-cultural study including more than 65,000 participants, showing that an increased mentalizing ability was indeed positively associated with supernatural beliefs (Maij et al., 2017). Thus, the hypothesized relation between mentalizing and supernatural beliefs appears robust, even though reported effect sizes are small and several studies have highlighted that despite mentalizing deficits, people scoring high on the autism quotient can still endorse supernatural beliefs and have supernatural encounters (Schaap-Jonker et al., 2013; Visuri, 2020).

Dualism

According to the naturalness of religion hypothesis (Bloom, 2007), humans have an early developing tendency to reason dualistically about the mind and the body. This tendency may be deeply engrained in our brain as we appear to have separate brain networks involved in reasoning about mental states (i.e., the theory-of-mind network and the default-mode-network) and for engaging in bodily processing (i.e., the fronto-parietal attention network; cf., Milliere, 2017). The bias for mind-body dualism already becomes prevalent from an early age onwards (Bering and Bjorklund, 2004; Bering et al., 2005): young children have a predisposition for applying dualistic reasoning about the mind and the body as being two separate entities, which might be at the basis of afterlife beliefs. In a large-scale cross-cultural study (using data collected in 24 countries across all 6 continents and including more than 10,000 participants) we set out to test the apparent cross-cultural universality of dualistic thinking (Hoogeveen and van Elk, submitted). To this end we presented participants with a vignette (using a similar design

as: Giménez and Harris, 2005) describing a grandmother who passed away, and we asked participants to make continuity judgments about physical (e.g., "Do you think she can still be hungry?") and mental states (e.g., "Do you think she can still love Bill?"). Overall, we found evidence for the hypothesis that the tendency to make continuity judgments for mental compared to physical states was cross-culturally prevalent, as participants judged mental states to be more likely to continue to exist than mental states. However, at the same time most people indicated cessation rather than continuation for all states (i.e., the modal response was to indicate that both mental and physical states would cease to exist after a person died), calling into question the apparent universality of mind-body dualism. Instead, the data appear more in line with an intuitive materialism account (Barrett et al., 2021), according to which the default is to view death in biological terms upon which all mental activity ends.

Intuitive thinking

Dual-process accounts of religion suggest that supernatural beliefs are primarily related to an intuitive (compared to an analytical) thinking style, whereas disbelief is related to analytical thinking (Pennycook et al., 2012). In other words: believers may be more prone to accept intuitive ideas and may have a reduced tendency for detecting cognitive conflict between potentially contradictory beliefs. An initial study attempted to show that priming analytical thinking reduces supernatural beliefs (Gervais and Norenzayan, 2012), however this finding could not be replicated in a high-powered replication study (Sanchez et al., 2017). In a large-scale cross-cultural study moreover, we found that the hypothesized relationship between religiosity and intuitive thinking was cross-culturally highly variable and only became apparent in three out of the 13 countries that were included (Gervais and Norenzayan, 2018). Other research also calls into question the presumed generic relationship between conflict detection and religiosity. For instance, in a registered report fMRI study we failed to find evidence for a negative relationship between religiosity and neural conflict responses (i.e., activity in the anterior cingulate cortex in response to a Stroop-task; cf., Hoogeveen et al., 2020). Other labs have shown similar null-results when attempting to replicate the relation between religiosity and intuitive thinking (Farias et al., 2017). The lack of a consistent relationship between intuitive thinking, conflict detection and supernatural beliefs could well be related to the lack of ecologically valid measures. For instance, the cognitive reflection task—one of the most widely used measures to assess analytical thinking—has been criticized for conflating mental abilities with processes (Blacksmith et al., 2019) and it is questionable whether making errors on a Stroop task relates in any meaningful way to the anxiety-relieving effects of religion.

In sum, there appears to be mixed evidence for the role of mentalizing, dualistic reasoning and intuitive thinking as cognitive precursors underlying supernatural beliefs.

Psychological functions subserved by supernatural beliefs

It has often been suggested that religion and supernatural beliefs can provide a palliative mechanism for coping with stressful events (Inzlicht et al., 2011), resonating with Karl Marx's adage that religion is opium for the people. Specifically, it has been suggested that religion helps us to cope with a lack of control and can provide direct benefits for one's mental and physical health.

Control

According to compensatory control theory (CCT), belief in a controlling God provides a palliative mechanism to cope with a lack of control (Kay et al., 2010). This theory is supported by a large amount of experimental findings showing that inducing a control threat manipulation (e.g., thinking back about a situation in which they lacked control) increased a compensatory efforts for restoring one's sense of control, such as an increased tendency to see illusory patterns (Whitson and Galinsky, 2008) and a preference for stage compared to continuous theories of development and evolution (Rutjens et al., 2013). However, in a registered report study (Hoogeveen et al., 2018) we failed to find evidence for an effect of lack of control on increased belief in a controlling God. However, we found—again in line with CCT—that in the US (but not in the Netherlands), experiencing less control in one's life in general, was associated with an increased belief in a controlling God.

Health

A wealth of studies have shown the positive effects of believing in God, religious practices (e.g., prayer and church attendance) and religious experiences on feelings of control, mental health and wellbeing (see for instance: Braam and Koenig, 2019; Garssen et al., 2021). However, most of these studies have been conducted in highly religious countries, thereby calling into question the cross-cultural generalizability of these findings. In a large-scale cross-cultural study, involving data collected in 24 countries across 5 continents, we set out to determine the boundary conditions of the religion-health relationship (Hoogeveen et al., 2022b). We used a many-analyst approach, whereby the data analysis was outsourced to 120 analysis teams who independently analyzed the data. Synthesizing the findings from these teams provided strong

evidence for the hypothesis that (1) religiosity is indeed positively associated with increased mental and physical wellbeing and (2) that this relationship depends on the perceived cultural norms of religiosity. Specifically: in highly religious countries such as the US or India, being religious is beneficial for one's health, whereas in more secular countries this relationship is absent or even reversed.

In sum, religiosity appears to have a positive relationship with mental health and can provide a sense of control, but only in countries in which being religious is the social norm.

Socio-cultural mechanisms underlying supernatural beliefs

Cultural-evolutionary accounts of religiosity have pointed out the relevance of socio-cultural factors for the learning and proliferation of supernatural beliefs (Norenzayan and Gervais, 2013), including religious role models and source heuristics.

Credibility-enhancing displays

CREDS are ostensible markers of religious commitments such as visiting religious services, wearing religious clothing, or adhering to a specific diet. CREDS have been suggested to be a strong predictor of the extent to which supernatural beliefs are transmitted from parents to children, as important role models do not just "talk the talk, but also walk the walk" (Henrich, 2009). This finding fits in a broader literature proposing that ultimately supernatural beliefs subserve an adaptive function by fostering in-group cohesion, cooperation and prosocial behavior (Norenzayan et al., 2016). Indeed, in a cross-cultural study we found that CREDS displayed by one's parents, were the strongest predictor of supernatural beliefs—much more so compared to thinking style, agency detection or mentalizing abilities (Maij et al., 2017). Thus, central role models during one's development, have a strong impact on the proliferation of supernatural beliefs.

Source heuristics

Next to CREDs, in general people appear more willing to trust information from sources that they credit with authority. The so-called Guru-effect refers to the observation that incomprehensible statements originating from a Guru are perceived to be meaningful, thereby only adding to the status of the Guru (Sperber, 2010). By using a vignette study in which participants were presented with seemingly profound statements that were attributed to a Guru or to a scientist, we found what we dubbed the *Einstein-effect*: across the globe participants rated the statement from the scientist as more profound than from the Guru (Hoogeveen et al., 2022a). We also found that this effect interacted with one's worldview: the Einstein-effect was most pronounced for atheist participants, but religious participants tended to attribute significance to statements from both the scientist and the guru (Hoogeveen et al., 2022a; van der Miesen et al., 2022). Source heuristics provide a proximate mechanism underlying the transmission of supernatural beliefs, and through their down-stream effects on cognitive processing (i.e., the down-regulation of executive functioning; Schjoedt et al., 2011) they also directly underlie the induction of placebo- and expectancy effects.

In sum, we found strong evidence for the role of CREDs and source heuristic effects in the proliferation and acceptance of supernatural beliefs.

Discussion

Why do some people believe in supernatural phenomena, whereas others don't? The research reviewed in this opinion paper points to the central relevance of socio-cultural factors for acquiring and maintaining supernatural beliefs. Rather than being rooted in deeply engrained tendencies for agency detection, mentalizing, reduced conflict detection or dualistic reasoning, the available evidence points toward the role of cultural scaffolding and explicit teaching for endorsing supernatural beliefs. Children are more likely to endorse the faith of their parents in case their parents engaged in ostensible religious displays. And in general people appear more willing to attribute significance to information from a source they deem trustworthy. Once these supernatural beliefs have been acquired, they encourage a self-sustaining loop by fostering agency-detection experiences, dualistic thinking, and encouraging a more intuitive processing style, providing a feeling of control and even having a protective effect on one's mental and physical health.

Author contributions

The author confirms being the sole contributor of this work and has approved it for publication.

Funding

This study was funded by Dr. Rüdiger Seitz, *via* the Volkswagen Foundation, Siemens Healthineers, and the Betz Foundation.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Barrett, H. C., Bolyanatz, A., Broesch, T., Cohen, E., Froerer, P., Kanovsky, M., et al. (2021). Intuitive dualism and afterlife beliefs: a cross-cultural study. *Cogn. Sci.* 45, e12992. doi: 10.1111/cogs.12992

Barrett, J. L. (2000). Exploring the natural foundations of religion. *Trends Cogn. Sci.* 4, 29–34. doi: 10.1016/S1364-6613(99)01419-9

Bering, J. M., and Bjorklund, D. F. (2004). The natural emergence of reasoning about the afterlife as a developmental regularity. *Dev. Psychol.* 40, 217–233. doi: 10.1037/0012-1649.40.2.217

Bering, J. M., Blasi, C. H., and Bjorklund, D. F. (2005). The development of afterlife beliefs in religiously and secularly schooled children. *Br. J. Dev. Psychol.* 23, 587–607. doi: 10.1348/026151005X36498

Biello, D. (2007). Searching for god in the brain. *Sci. Am. Mind* 185, 38–45. doi: 10.1038/scientificamericanmind1007-38

Blacksmith, N., Yang, Y., Behrend, T. S., and Ruark, G. A. (2019). Assessing the validity of inferences from scores on the cognitive reflection test. *J. Behav. Dec. Making* 32, 599–612. doi: 10.1002/bdm.2133

Bloom, P. (2007). Religion is natural. *Dev. Sci.* 10, 147–151. doi: 10.1111/j.1467-7687.2007.00577.x

Boyer, P. (2001). *Religion Explained: The Evolutionary Origins of Religious Thought*. New York: Basic Books.

Braam, A. W., and Koenig, H. G. (2019). Religion, spirituality and depression in prospective studies: a systematic review. *J. Affect Disord.* 257, 428–438. doi: 10.1016/j.jad.2019.06.063

Charles, S. J., Bartlett, J. E., Messick, K. J., Coleman, I. I. I., T. J., and Uzdavines, A. (2019). Researcher degrees of freedom in the psychology of religion. *Int. J. Psychol. Relig.* 29, 230–245. doi: 10.1080/10508619.2019.1660573

Farias, M., van Mulukom, V., Kahane, G., Kreplin, U., Joyce, A., Soares, P., et al. (2017). Supernatural belief is not modulated by intuitive thinking style or cognitive inhibition. *Sci. Rep.* 7, 15100. doi: 10.1038/s41598-017-14090-9

Garssen, B., Visser, A., and Pool, G. (2021). Does spirituality or religion positively affect mental health? Meta-analysis of longitudinal studies. *Int. J. Psychol. Relig.* 31, 4–20. doi: 10.1080/10508619.2020.1729570

Gervais, W. M., and Norenzayan, A. (2012). Analytic thinking promotes religious disbelief. *Science* 336, 493–496. doi: 10.1126/science.1215647

Gervais, W. M., and Norenzayan, A. (2018). Analytic atheism revisited. *Nat. Hum. Behav.* 2, 609. doi: 10.1038/s41562-018-0426-0

Giménez, M., and Harris, P. (2005). Children's acceptance of conflicting testimony: the case of death. *J. Cogn. Cult.* 5, 143–164. doi: 10.1163/1568537054068606

Henrich, J. (2009). The evolution of costly displays, cooperation and religion: credibility enhancing displays and their implications for cultural evolution. *Evol. Hum. Behav.* 30, 244–260. doi: 10.1016/j.evolhumbehav.2009.03.005

Hoogeveen, S., Haaf, J. M., Bulbulia, J. A., Ross, R. M., McKay, R., Altay, S., et al. (2022a). The Einstein effect provides global evidence for scientific source credibility effects and the influence of religiosity. *Nat. Hum. Behav.* 6, 523–535. doi: 10.1038/s41562-021-01273-8

Hoogeveen, S., Sarafoglou, A., van Elk, M., and Wagenmakers, E. J. (2022b). A many-analysts approach to the relation between religiosity and well-being: the dataset. *Relig. Brain Behav*. doi: 10.1080/2153599X.2022.2070255

Hoogeveen, S., Snoek, L., and van Elk, M. (2020). Religious belief and cognitive conflict sensitivity: a preregistered fMRI study. *Cortex* 129, 247–265. doi: 10.1016/j.cortex.2020.04.011

Hoogeveen, S., and van Elk, M. (2018). Advancing the cognitive science of religion through replication and open science. *J. Cogn. Sci. Relig.* 6, 158–190. doi: 10.1558/jcsr.39039

Hoogeveen, S., and van Elk, M. (submitted). *A Cross-Cultural Investigation of Mind-Body Dualism and Religiosity*. Cognition.

Hoogeveen, S., Wagenmakers, E. J., Kay, A. C., and van Elk, M. (2018). Compensatory control and religious beliefs: A registered replication report across two countries. *Compr. Results Soc. Psychol.* 3, 240– 65.

Inzlicht, M., Tullett, A. M., and Good, M. (2011). The need to believe: a neuroscience account of religion as a motivated process. *Relig. Brain Behav.* 1, 192–212. doi: 10.1080/2153599X.2011.647849

Kay, A. C., Gaucher, D., McGregor, I., and Nash, K. (2010). Religious belief as compensatory control. *Pers. Soc. Psychol. Rev.* 14, 37–48. doi: 10.1177/1088868309353750

Maij, D. L. R., van Harreveld, F., Gervais, W., Schrag, Y., Mohr, C., and van Elk, M. (2017). Mentalizing skills do not differentiate believers from non-believers, but credibility enhancing displays do. *PLoS ONE* 12, e0182764. doi: 10.1371/journal.pone.0182764

Milliere, R. (2017). Looking for the self: phenomenology, neurophysiology and philosophical significance of drug-induced ego dissolution. *Front. Hum. Neurosci.* 11, 245. doi: 10.3389/fnhum.2017.00245

Norenzayan, A., and Gervais, W. M. (2013). The origins of religious disbelief. *Trends Cogn. Sci.* 17, 20–25. doi: 10.1016/j.tics.2012.11.006

Norenzayan, A., Gervais, W. M., and Trzesniewski, K. H. (2012). Mentalizing deficits constrain belief in a personal God. *PLoS ONE* 7, e36880. doi: 10.1371/journal.pone.0036880

Norenzayan, A., Shariff, A. F., Gervais, W. M., Willard, A. K., McNamara, R. A., Slingerland, E., et al. (2016). The cultural evolution of prosocial religions. *Behav. Brain Sci* 39, e1. doi: 10.1017/S0140525X14001356

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., et al. (2015). Scientific standards. Promoting an open research culture. *Science* 348, 1422–1425. doi: 10.1126/science.aab2374

Pennycook, G., Cheyne, J. A., Seli, P., Koehler, D. J., and Fugelsang, J. A. (2012). Analytic cognitive style predicts religious and paranormal belief. *Cognition* 123, 335–346. doi: 10.1016/j.cognition.2012.03.003

Rutjens, B. T., van Harreveld, F., van der Pligt, J., Kreemers, L. M., and Noordewier, M. K. (2013). Steps, stages, and structure: finding compensatory order in scientific theories. *J. Exp. Psychol. Gen.* 142, 313–318. doi: 10.1037/a0028716

Sanchez, C., Sundermeier, B., Gray, K., and Calin-Jageman, R. J. (2017). Direct replication of Gervais and Norenzayan (2012): No evidence that analytic thinking decreases religious belief. *PLoS ONE* 12, e0172636. doi: 10.1371/journal.pone.0172636

Schaap-Jonker, H., Sizoo, B., Van Schothorst-Van Roekel, J., and Corveleyn, J. (2013). Autism spectrum disorders and the image of God as a core aspect of religiousness. *Int. J. Psychol. Relig.* 23, 145–160. doi: 10.1080/10508619.2012.688005

Schjoedt, U., Stodkilde-Jorgensen, H., Geertz, A. W., Lund, T. E., and Roepstorff, A. (2011). The power of charisma–perceived charisma inhibits the frontal executive network of believers in intercessory prayer. *Soc. Cogn. Affect Neurosci.* 6, 119–127. doi: 10.1093/scan/nsq023

Schjoedt, U., Stodkilde-Jorgensen, H., Geertz, A. W., and Roepstorff, A. (2009). Highly religious participants recruit areas of social cognition in personal prayer. *Soc. Cogn. Affect Neurosci.* 4, 199–207. doi: 10.1093/scan/nsn050

Sperber, D. (2010). The guru effect. *Rev. Philos. Psychol.* 1, 583–592. doi: 10.1007/s13164-010-0025-0

Vail, K. E., III, Rothschild, Z. K., Weise, D. R., Solomon, S., Pyszczynski, T., and Greenberg, J. (2010). A terror management analysis of the psychological functions of religion. *Pers. Soc. Psychol. Rev.* 14, 84–94. doi: 10.1177/1088868309351165

van der Miesen, M. M., van der Lande, G. J., Hoogeveen, S., Schjoedt, U., and van Elk, M. (2022). The effect of source credibility on the evaluation of statements in a spiritual and scientific context: a registered report study. *Compr. Results Soc. Psychol.* 1–26. doi: 10.1080/23743603.2022.2041984

van Elk, M. (2013). Paranormal believers are more prone to illusory agency detection than skeptics. *Conscious. Cognit.* 22, 1041–1046. doi: 10.1016/j.concog.2013.07.004

van Elk, M. (2015). Perceptual biases in relation to paranormal and conspiracy beliefs. *PLoS ONE* 10, e0130422. doi: 10.1371/journal.pone.0130422

van Elk, M., Matzke, D., Gronau, Q. F., Guan, M., Vandekerckhove, J., and Wagenmakers, E. J. (2015). Meta-analyses are no substitute for registered replications: a skeptical perspective on religious priming. *Front. Psychol.* 6, 01365. doi: 10.3389/fpsyg.2015.01365

Visuri, I. (2020). Sensory supernatural experiences in autism. *Relig. Brain Behav.* 10, 151–165. doi: 10.1080/2153599X.2018.1548374

Whitson, J. A., and Galinsky, A. D. (2008). Lacking control increases illusory pattern perception. *Science* 322, 115–117. doi: 10.1126/science.1159845

Willard, A. K., and Norenzayan, A. (2013). Cognitive biases explain religious belief, paranormal belief, and belief in life's purpose. *Cognition* 129, 379–391. doi: 10.1016/j.cognition.2013.07.016

frontiers | Frontiers in Psychology

TYPE Mini Review
PUBLISHED 03 August 2022
DOI 10.3389/fpsyg.2022.938731

Check for updates

OPEN ACCESS

EDITED BY
Rüdiger J. Seitz,
Heinrich Heine University of Düsseldorf, Germany

REVIEWED BY
Piotr Balcerowicz,
University of Warsaw, Poland

*CORRESPONDENCE
Jed Forman
jed.forman@simpson.edu

SPECIALTY SECTION
This article was submitted to Theoretical and Philosophical Psychology, a section of the journal Frontiers in Psychology

RECEIVED 08 May 2022
ACCEPTED 29 June 2022
PUBLISHED 03 August 2022

CITATION
Forman J (2022) Believing is seeing: A Buddhist theory of creditions. *Front. Psychol.* 13:938731. doi: 10.3389/fpsyg.2022.938731

COPYRIGHT
© 2022 Forman. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Believing is seeing: A Buddhist theory of creditions

Jed Forman*

Center for Buddhist Studies, University of California, Berkeley, Berkeley, CA, United States

The creditions model is incredibly powerful at explaining both how beliefs are formed and how they influence our perceptions. The model contains several cognitive loops, where beliefs not only influence conscious interpretations of perceptions downstream but are active in the subconscious construction of perceptions out of sensory information upstream. This paper shows how this model is mirrored in the epistemology of two central Buddhist figures, Dignāga (480–540 CE) and Dharmakīrti (c. 550–650 CE). In addition to showing these parallels, the paper also demonstrates that by drawing on Dignāga and Dharmakīrti's theory, we can extend the explanatory power of the creditions model. Namely, while creditions explain how beliefs influence both the conscious interpretation and subconscious construction of sensory information, Dignāga and Dharmakīrti suggest beliefs can even be *generative* of sensory-like information. I recruit ancient Buddhist texts in conjunction with contemporary cognitive science scholarship to offer a hypothesis for the cognitive mechanisms responsible for this.

KEYWORDS

perception, epistemology, Buddhism, cognitive science, meditation

Dignāga and Dharmakīrti's epistemology

Dignāga and Dharmakīrti's epistemology advocates a sharp divide between perception and inference. On their view, perception is our direct encounter with the world, namely (though not exclusively, as we will see) through the senses. They understand perception as largely causal, with external objects affecting the senses to produce a perception. Inference, on the other hand, uses perceptual information to adduce non-perceptual facts.

The classic example of inference is that of fire from smoke. Because smoke is necessarily created by fire, the perception of smoke warrants an inference of fire. Thus, even when a fire is occluded from our sight, one is justified in concluding there is a fire present after seeing smoke rising. Dignāga and Dharmakīrti argue that these two epistemic instruments (perception and inference) give an exhaustive epistemology, explaining all instances of warranted knowledge.

Dignāga and Dharmakīrti's differentiation between perception and inference has led some authors to conclude that their theory is a species of sense-data theory. That is,

while we perceive colors, shapes, sounds, or textures, we use this information to infer the presence of common-sense objects and medium-sized dry goods. On this view, one never even directly perceives smoke. Instead, one perceives gray forms that are inferred to be "smoke," and based on this inference, one further infers fire (Arnold, 2017, para. 24, Arnold, 2019, p. 227–228). If this were Dignāga and Dharmakīrti's position, it would pit them close to the philosophy of Alfred J. Ayer, who argued that common sense objects are inferred based upon our perception of sense data (Ayer, 1967, p. 129).

There are some aspects of Dignāga's and Dharmakīrti's thought that suggest a sense-data theory. Like Ayer, they do argue that we do not perceive medium-sized dry goods, like tables, chairs, peoples, and trees. They consider such objects to be merely conceptual (*vikalpaka*) constructs, reified "universals" (*sāmānya*). As such, they are the referent objects of inferences. Reality itself, on the other hand, is composed of discrete particles that only last for a moment. These are particulars (*svalakṣaṇa*). On this theory, we could think of reality like a buzzing soup of static and white noise. Our tendency to construe enduring, extended objects out of this soup is like a case of ongoing apophenia, the recognition of patterns in otherwise random data.

However, Dignāga and Dharmakīrti's theory of conceptualization is distinct from apophenia in an important regard. Unlike apophenia, concepts have pragmatic utility (*arthakriyā*). Dharmakīrti gives an analogy to a jewel to make this point. Two people see some shimmering light, and both think that it is a jewel reflecting light. Both cognitions are erroneous (*bhrānti*), since (according to Buddhists) no universal "jewel" inheres in the world. Nevertheless, in one case the light is produced by a lamp and in the other by a group of particulars that collectively have the qualities we would expect of a jewel. In the latter case, then, the cognition is informative (*saṃvāda*) despite being erroneous, since we can use that cognition to reach particulars that behave in the way we expect of a jewel, even if no jewel is there *really* (Miyasaka, 1972, 2:v.3.57-8; Devendrabuddhi, 1744, F. 145a−146b).

In some ways, this is compatible with Ayer's (perhaps counterintuitive) notion of inference. As a logical positivist, Ayer agrees that the ultimate arbiter of our cognitions is their efficacy, and not whether they represent "real" things. Nevertheless, Dignāga and Dharmakīrti's theory of conceptual construction does *not* entail sense data theory. This is because they consider even the apprehension of color to be a conceptual process, a construction of a universal. As Dignāga states, "The apprehension of a color, or the like, [arises] from both the particular, which is ineffable (*avyapadeśya*), and a color, which is a universal" (Hattori, 1968 p. 24 and 81n1.19). In other words, even the recognition of some color involves a constructive process. This follows from Buddhist ontology, since even patches of color (no matter how small) are things that appear to take up time and space.

What, then, *is* perceived according to Dignāga and Dharmakīrti? They argue perception perceives particulars. Yet, as Dignāga states, because particulars are completely unique and momentary, they are "ineffable" (*avyapadeśya*). Thus, we cannot say anything about perceptual content, since any such saying is conceptual. This may seem mystical at first. But if we understand perception causally, it becomes less so. "Perception" just means the causal interaction between the senses and the world. It has no content to speak of. Such content only arises to awareness once conceptual processes have done their work (see Sharf, 2018 for details).

This theory comes close to that of another thinker, Charles Peirce.[1] Like Dignāga and Dharmakīrti, Peirce also argues perception is "subconscious" and not operative at the level of awareness. In place of Ayer's inference, he appeals to "abduction" to bridge the divide between perception and our awareness of medium-sized dry goods. Abduction involves pragmatic heuristics that help us navigate our world even though they may misrepresent reality. They are thus "extremely fallible" and updatable as new information arises (Peirce, 1955, p. 304). Like Peirce, Dharmakīrti argues that our conceptualizations do not have any necessary authenticity, but are the product of certain "patterns of thought" (*āhitā vāsanā*) (Gnoli, 1960, 42 ll.13–14). These patterns of thought are preserved or culled to the degree they help us get what we want and avoid what we do not want (Mikogami, 1979).[2]

Believing is seeing

For anyone familiar with the creditions model of belief formation, Dignāga and Dharmakīrti's theory will appear familiar. In the creditions model, awareness of perceptual information only comes at the end of a multistep process. Such information is first parsed through pre-linguistic, "primal" beliefs that are predictive. Like in Dharmakīrti's jewel analogy, such beliefs might predict finding a jewel based on the perception of shimmering light. Also like in Dharmakīrti's analogy, these beliefs can be refined based on their efficacy. So, if someone sees a shimmer but does not find a jewel, such shimmers will be less likely to produce the assumption of a jewel in the future.

Rüdiger Seitz describes two ways in which these primal beliefs can be updated. The first is through the processing of prediction errors. The person who does not find a jewel updates

1 Dunne (2004, p. 49) also notes parallels between Dharmakīrti and Peirce, though he cautions not to overstate the similarities.

2 In Dharmakīrti's epistemology, "conceptual habituation" (*vikalpābhyāsa*) probably comes closest to Peirce's abduction, which describes how repeated experience leads to an automatic cognitive association between the perceptual stimulus and a given concept (Eltschinger, 2014, §1.2; Kellner, 2004, p. 30–31).

their valuation processes spontaneously so that they make better predictions. This occurs below the level of awareness. However, these processes can also be updated via conscious awareness. Because beliefs can be semantically encoded into language, we can become aware of them. By reflecting on these beliefs, the brain can affect valuations, changing beliefs and the processing of perceptual information (Seitz et al., 2019; Seitz and Angel, 2020; Seitz, 2022a,b). For example, by reflecting on the irrationality of racist beliefs, one can affect their snap judgements about others.

The creditions model is thus abductive in Peirce's sense and pragmatic in Dharmakīrti's. All three models understand cognitive processing to be fallibilist rather than apodictic, updating itself as information arises. However, both Peirce and Seitz present these updates as a transformation of the valuation process. In other words, while the flow of perceptual information stays consistent, it is only how the information is processed that is affected. It is on this point that Dharmakīrti offers a variant theory.

On Dharmakīrti's theory, perception has greater epistemic weight than inference. This is because all inference is erroneous. To comprehend Buddhist ideas deeply, then, Dharmakīrti argues the practitioner must *perceive* these truths in addition to understanding them conceptually. This perceptual understanding is achieved not by sensory perception, but by a special type of perception called "yogic perception" (*yogipratyakṣa*). Dharmakīrti explains yogic perception is the product of sustained meditation. He claims that by meditating on some universal, holding it in the mind's eye, the meditator will eventually have "a nonconceptual clear appearance constructed by the power of meditation." Although this is not an instance of sensory perception, Dharmakīrti argues that its clarity is qualitatively indistinguishable from "seeing" something "as if it were right in front of them" (Miyasaka, 1972, 2:v.3.282-4).

Admittedly, it is somewhat unclear what it would be like to "see" an abstract Buddhist concept in such a vivid manner. Nevertheless, Dharmakīrti presents an intriguing possibility. If we think of meditation as a type of reflection, Dharmakīrti argues that reflective processes do not just affect valuation systems, but perceptual systems as well. In other words, reflection might *generate* perceptual information, not merely *affect* how that information is processed.

In this regard, Dharmakīrti offers several analogies to cognitive processes similar to yogic perception. Specifically, he cites hallucinations that are caused by intense emotion, such was when "one is driven crazy by desire, fear, or grief" (Miyasaka, 1972, 2:v.3.282). Dharmakīrti's assertion that grief can lead to hallucinations is well documented. Indeed, vivid hallucinations of the deceased are not uncommon during bereavement (Castelnovo et al., 2015). Dharmakīrti argues that intense rumination on a loved one eventually spills over into a perceptual event, such that they are no longer just in the mind's eye but seen "as if they were right in front" of the griever. Meditation operates through the same mechanism. By fixating on an idea for a sustained period of time, it will eventually appear clearly and perceptually (Miyasaka, 1972, 2:3.285-6).

Cognitive underpinnings

Dharmakīrti wants to differentiate yogic perception from meditative hallucinations. It is only when the initial meditative idea is "true" that the resultant perception is yogic (Miyasaka, 1972, 2:3.286).[3] This epistemological issue aside, I want to focus on the mechanisms for how meditation might be generative of novel perceptual content, since the creditions model does not account for such a possibility, nor how it might influence belief formation.

For example, Seitz explains hallucinations as either misinterpretations "triggered by items in the patient's environment" or arising "spontaneously," perhaps as cognitive misfires (Seitz, 2022a, p. 27). Phillip Gerrans also understands hallucinations as false valuations of perceptual events, "an imaginative state triggered by a sensory or perceptual anomaly" (Gerrans, 2014, p. 137). Seitz's and Gerrans' model would theorize grief hallucinations as the product of over-interpreting sensory information, leading to the sensed presence of a missed loved one. Justin Barrett gives a similar account of the apparition of supernatural agents, where beliefs manipulate the interpretation of sensory information so that bumps and creaks in the night become confirmations of ghosts (Barrett, 2004, chap. 3).

While, indeed, many hallucinations are the product of misinterpretations, others appear too phenomenologically rich to be the result of exaggerations upon sparse perceptual data. For example, consider the following account of a man grieving the loss of his father. The man claims he "was certainly awake" and saw his deceased father in the middle of the night "sitting on the corner of my bed … He was opaque, not ethereal in any way." What is even more telling about this event is that the griever did not *believe* that he really saw his father. "I do not know whether this was a hallucination or something else, but since I provisionally do not believe in the paranormal, it must have been" (Sacks, 2012, chap. 13). In other words, the hallucination did not appear to be the result of a proclivity to over interpret sensory information to conform with preexisting beliefs. Rather, the hallucination had a perceptual richness *despite* his belief to the contrary. This suggests that something about the reflective process affects not just how perceptual information is interpreted, but can

3 There is debate in the secondary literature as to how Dharmakīrti make this differentiation. Compare Dunne (2007, p. 515) vs. Eltschinger (2009, 169n1) and Franco (2011, 87 ff.).

generate perceptual content, even when that content contradicts reflective beliefs.

Although this account is only anecdotal, there is a wealth of evidence that suggests hallucinations can originate from top-down processes, like rumination, in this fashion. To be sure, much, if not the majority, of hallucinatory phenomena is the result of some imbalance between bottom-up perceptual information and top-down predictive coding. Nevertheless, hallucinations can also be the result of top-down processes unilaterally affecting the visual cortex, such as the suppression of sensory signals by the prefrontal cortex (Ranson et al., 2019), coupling between the default mode network (DMN) and the visual cortex (Walpola et al., 2020), and visual cortex activation by higher cortical areas during visualization (Howe and Carter, 2016). The last two examples are especially pertinent to the case of meditation, since what Dharmakīrti has in mind is an intense visualization practice—which is either instigated be intense emotion, such as grief, or the result of deliberate cultivation. Several studies reveal that meditation increases DMN-visual-cortex coupling (Faber et al., 2014; Berkovich-Ohana et al., 2016; Fujino et al., 2018; Zhang et al., 2021), which may offer a mechanism of how deliberate meditation induces hallucinations.

Another possible mechanism to explain vivid hallucinations induced by meditation is hypnosis. Some research suggests that hypnosis and meditation create vivid visual experiences through a shared mechanism. Namely, both *downregulate* executive prefrontal systems as well as the DMN (Dietrich and Al-Shawaf, 2018), creating a space within which imaging systems can create vivid representations from the bottom-up, unimpeded by prefrontal regulation (Winkelman, 2017). Even though meditation is highly focused, the high recruitment of attentional systems in both meditation and hypnosis creates hypoactivity in other prefrontal systems, leading to deregulation (Dietrich and Al-Shawaf, 2018). However, recent scholarship has brought this hypofrontality thesis into question (Fingelkurts et al., 2007; Facco, 2021). Thus, other scholarship concludes that hypnosis enhances the vividness of mental imagery top-down via the prefrontal cortex (Sireteanu et al., 2010; Lanfranco et al., 2021). This might explain how images in the mind's eye can become vividly visual *via* deliberate meditative practice.

In sum, this research suggests at least three possible mechanisms through which meditation might produce perceptual content: (1) the coupling of the visual cortex with other cognitive systems, (2) the downregulation of prefrontal systems, letting imagery bubble up from the bottom up, and (3) the creation of vivid imagery from the top down. It is not unlikely that all these procedural alternatives are possible, meaning that visual hallucination is overdetermined by meditative practice. Indeed, there are many different types of meditative practices, each of which may exploit these pathways differently.

Our analysis thus reveals that higher-order cognitive processes, like reflection, might not just transform how perceptual information is processed, but may generate perceptual content itself. In other words, belief may not just manipulate *how* we see but generate *what* we see.

Conclusion

Dignāga and Dharmakīrti's theory thus shares many affinities with the creditions model, particularly concerning how belief formation develops under normal circumstances. Both theories argue that perception is causal and subconscious, that perceptual awareness is highly entangled with beliefs about the world, and that these beliefs are fallible, formed by abductive processes that are patterned by experience.

Dignāga and Dharmakīrti, however, present an additional picture of how this processing can flow in special circumstances. That is, beliefs do not just organize perceptual information upstream nor merely interpret that information downstream. In rare cases, beliefs can generate perceptual information itself. Dignāga and Dharmakīrti argue that it is only in some cases that this process is epistemic, when these starting beliefs are "true." But if we bracket epistemology, meditative hallucinations may be instrumental in *belief* formation, regardless of whether those beliefs constitute knowledge. For example, fixation on concepts like "ghost" might not just cause someone to *interpret* perceptual data as ghosts, but *produce* the perception of a ghost, reinforcing their belief in ghosts.

The hypothesis at this point is speculative. Future research could use fMRI imaging to gain a closer look at how meditation affects the visual cortex, and whether that activity is highly correlated with visual hallucinations. Such research should be sensitive to the meditative practices involved, particularly whether they are the type of concentration-demanding practices described by Dharmakīrti.

If meditation does prove to be generative of perceptual content in the fashion hypothesized, then it offers another important clue into the phenomenon of belief formation, especially of the religious sort. That is, religious beliefs might not merely arise as ways to make sense of aberrant sensory experiences (as in Seitz's and Gerrans' model), nor do they merely persist as intuitive explanations of our sensory world (Sperber, 1996, p. 98–118). In addition to these modes, beliefs may also generate their own perceptual content in a manner that makes them self-confirming. Tanya Luhrman's work has also explored this possibility (Luhrmann, 2012). If this is true, what remains to be seen is the pervasiveness of these experiences—whether they are only the provenance of elite practitioners engaged in meditative practices, or they are operable even

among a wider population and explain the persistence of their religious beliefs.

Author contributions

The author confirms being the sole contributor of this work and has approved it for publication.

Funding

This paper is funded by Dr. Rüdiger Seitz, *via* the Volkswagen Foundation, Siemens Healthineers, and the Betz Foundation. Siemens Healthineers were not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Arnold, D. (2017). "The philosophical works and influence of Dignāga and Dharmakīrti," in *Oxford Research Encyclopedia of Religion* (Oxford: Oxford University Press). doi: 10.1093/acrefore/9780199340378.013.198

Arnold, D. (2019). "Givenness as a corollary to non-conceptual awareness: thinking about thought in Buddhist philosophy," in *Wilfrid Sellars and Buddhist Philosophy: Freedom from Foundations, Google Play Edition* (New York: Routledge), 214–256.

Ayer, A. J. (1967). Has Austin refuted the Sense-Datum theory? *Synthese* 17, 117–140.

Barrett, J. L. (2004). *Why Would Anyone Believe in God?* Cognitive Science of Religion Series. Walnut Creek, CA: AltaMira Press.

Berkovich-Ohana, A., Harel, M., Hahamy, A., Arieli, A., and Malach, R. (2016). Alterations in task-induced activity and resting-state fluctuations in visual and DMN areas revealed in long-term meditators. *NeuroImage* 135, 125–134. doi: 10.1016/j.neuroimage.2016.04.024

Castelnovo, A., Cavallotti, S., Gambini, O., and D'Agostino, A. (2015). Post-bereavement hallucinatory experiences: a critical overview of population and clinical studies. *J. Affect. Disord.* 186, 266–274. doi: 10.1016/j.jad.2015.07.032

Devendrabuddhi. (1744). "*Pramāṇavārttikapañjikā, Tshad Ma Rnam 'grel Kyi Dka' 'grel," in *Sde Dge Bstan 'gyur*, edited by Tshul khrims rin chen, translated by Subhutiśri and Dge ba'i blo gros, Toh. no. 4217, Tshad ma, che:1b—326b. Delhi: Delhi Karmapae Choedhey, Gyalwae Sungrab Partun Khang.

Dietrich, A., and Al-Shawaf, L. (2018). The transient hypofrontality theory of altered states of consciousness. *J. Conscious. Stud.* 25, 226–247.

Dunne, J. D. (2004). *Foundations of Dharmakīrti's Philosophy*, *1st edn.* Boston: Wisdom Publications.

Dunne, J. D. (2007). Realizing the unreal: Dharmakīrti's theory of yogic perception. *J. Indian Philos.* 34, 497–519. doi: 10.1007/s10781-006-9008-y

Eltschinger, V. (2009). "On the Career and the Cognition of Yogins," in *Yogic Perception, Meditation and Altered States of Consciousness,* eds Franco, E., and Eigner, D. (Vienna: Verlag der Österreichischen Akademie der Wissenschaften), 169–213.

Eltschinger, V. (2014). The four nobles' truths and their 16 aspects: on the dogmatic and soteriological presuppositions of the Buddhist epistemologists' views on Niścaya. *J. Indian Philos.* 42, 249–273. doi: 10.1007/s10781-013-9193-4

Faber, P. L., Lehmann, D., Gianotti, L. R. R., Milz, P., Pascual-Marqui, R. D., Held, M., et al. (2014). Zazen meditation and no-task resting EEG compared with LORETA intracortical source localization. *Cogn. Process.* 16, 87–96. doi: 10.1007/s10339-014-0637-x

Facco, E. (2021). Hypnosis and hypnotic ability between old beliefs and new evidences: an epistemological reflection. *Am. J. Clin. Hyp.* 64, 20–35. doi: 10.1080/00029157.2020.1863181

Fingelkurts, A. A., Kallio, S., and Revonsuo, A. (2007). Cortex functional connectivity as a neurophysiological correlate of hypnosis: an EEG case study. *Neuropsychologia* 45, 1452–1662. doi: 10.1016/j.neuropsychologia.2006.11.018

Franco, E. (2011). "Perception of yogis—some epistemological and metapyhysical considerations," in *Religion and Logic in Buddhist Philosophical Analysis: Proceedings of the Fourth International Dharmakīrti Conference, Vienna*, August 23–27, 2005, eds Krasser, H., Lasic, H., Franco, E., and Kellner, B. (Beiträge Zur Kultur- Und Geistesgeschichte Asiens, Nr. 69. Wien: Verlag der Österreichischen Akademie der Wissenschaften), 81–98.

Fujino, M., Ueda, Y., Mizuhara, H., Saiki, J., and Nomura, M. (2018). Open monitoring meditation reduces the involvement of brain regions related to memory function. *Sci. Rep.* 8, 9968–9910. doi: 10.1038/s41598-018-28274-4

Gerrans, P. (2014). *The Measure of Madness: Philosophy of Mind, Cognitive Neuroscience, and Delusional Thought.* Cambridge, MA: The MIT Press.

Gnoli, R. (ed.) (1960). *The Pramanavarttikam of Dharmakīrti: The First Chapter with the Autocommentary*. Roma: Instituto italiano per il medio ed estremo oriente.

Hattori, M. (ed.) (1968). *Dignaga: On Perception, Being the Pratyakṣapariccheda of Dignaga's Pramāṇasamuccaya from the Sanskrit Fragments and the Tibetan Versions*. Translated by Masaaki Hattori. Harvard Oriental Series 47. Cambridge, MA: Harvard University Press.

Howe, P. D. L., and Carter, O. L. (2016). Hallucinations and mental imagery demonstrate top-down effects on visual perception. *Behav. Brain Sci.* 39, e248. doi: 10.1017/S0140525X15002502

Kellner, B. (2004). "Why infer and not just look? dharmak irti on the psychology of inferential processes," in *The Role of the Example (Dṛṣṭānta) in Classical Indian Logic*, eds. Katsura, S., and Steinkellner, E. (Vienna: Arbeitskreis fu r tibetische und buddhistische Studien Universita t Wien), 1–51.

Lanfranco, R. C., Rivera-Rei, Á., Huepe, D., Ibáñez, A., and Canales-Johnson, A. (2021). Beyond imagination: hypnotic visual hallucination induces greater lateralised brain activity than visual mental imagery. *NeuroImage* 239, 118282. doi: 10.1016/j.neuroimage.2021.118282

Luhrmann, T. M. (2012). *When God Talks Back: Understanding the American Evangelical Relationship with God*, *First edition*. New York: Alfred A. Knopf.

Mikogami, E. (1979). Some remarks on the concept of Arthakriyā. *J. Indian Philos.* 7, 79–94.

Miyasaka, Y. (ed) (1972). *Pramāṇavarttika-kārikā: (Sanskrit and Tibetan)*, *Vol. 2.* Acta Indologica. Narita: Naritasan Shinshoji.

Peirce, C. S. (1955). *Philosophical Writings of Peirce*. Edited by Justus Buchler. New York: Dover Publications.

Ranson, A., Broom, E., Powell, A., Chen, F., Major, G., and Hall, J. (2019). Top-down suppression of sensory cortex in an NMDAR hypofunction model of psychosis. *Schizophrenia Bull.* 45, 1349–1357. doi: 10.1093/schbul/sby190

Sacks, O. (2012). *Hallucinations*. New York: Knopf Doubleday Publishing Group.

Seitz, R. J. (2022a). Beliefs: a challenge in neuropsychological disorders. *J. Neuropsychol.* 16, 21–37. doi: 10.1111/jnp.12249

Seitz, R. J. (2022b). Believing and beliefs—neurophysiological underpinnings. *Front. Behav. Neurosci.* 16, 880504. doi: 10.3389/fnbeh.2022.880504

Seitz, R. J., and Angel, H.-F. (2020). Belief formation—a driving force for brain evolution. *Brain Cogn.* 140, 105548. doi: 10.1016/j.bandc.2020.105548

Seitz, R. J., Paloutzian, R. F., and Angel, H. (2019). Believing is representation mediated by the dopamine brain system. *Eur. J. Neurosci.* 49, 1212–1214. doi: 10.1111/ejn.14317

Sharf, R. H. (2018). Knowing blue: early buddhist accounts of non-conceptual sense. *Philos. East West* 68, 826–870. doi: 10.1353/pew.2018.0075

Sireteanu, R., Oertel, V., Mohr, H., Haenschel, C., Linden, D., Maurer, K., et al. (2010). Graphical illustration and functional neuroimaging of visual hallucinations during prolonged blindfolding: a comparison to visual imagery. *J. Vis.* 8, 1805–1822. doi: 10.1167/8.6.68

Sperber, D. (1996). *Explaining Culture: A Naturalistic Approach*. Malden, MA: Blackwell.

Walpola, I. C., Muller, A. J., Hall, J. M., Andrews-Hanna, J. R., Irish, M., Lewis, S. J. G., et al. (2020). Mind-wandering in Parkinson's disease hallucinations reflects primary visual and default network coupling. *Cortex* 125, 233–245. doi: 10.1016/j.cortex.2019.12.023

Winkelman, M. J. (2017). The mechanisms of psychedelic visionary experiences: hypotheses from evolutionary psychology. *Front. Neurosci.* 11, 539. doi: 10.3389/fnins.2017.00539

Zhang, Z., Luh, W. M., Duan, W., Zhou, G. D., Weinschenk, G., Anderson, A. K., et al. (2021). Longitudinal effects of meditation on brain resting-state functional connectivity. *Sci. Rep.* 11, 11361. doi: 10.1038/s41598-021-90729-y

frontiers | Frontiers in Psychology

TYPE Opinion
PUBLISHED 02 November 2022
DOI 10.3389/fpsyg.2022.1033021

Check for updates

OPEN ACCESS

EDITED BY
Rüdiger J. Seitz,
Heinrich Heine University of Düsseldorf, Germany

REVIEWED BY
Raymond Paloutzian,
Westmont College, United States

*CORRESPONDENCE
Li-Jie Du
weareglad@126.com

SPECIALTY SECTION
This article was submitted to Theoretical and Philosophical Psychology, a section of the journal Frontiers in Psychology

RECEIVED 31 August 2022
ACCEPTED 17 October 2022
PUBLISHED 02 November 2022

CITATION
Du L-J (2022) Zen and the psychological significance of meditation as related to believing. *Front. Psychol.* 13:1033021. doi: 10.3389/fpsyg.2022.1033021

COPYRIGHT
© 2022 Du. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Zen and the psychological significance of meditation as related to believing

Li-Jie Du*

Department of Sociology, College of Philosophy, Law, and Politics, Shanghai Normal University, Shanghai, China

KEYWORDS

sitting meditation, Zen Buddhism, psychological significance, mental health, Rinzai Zen, believing

Introduction

The word "Zen" in Buddhism is derived from the Sanskrit word "Dhyāna,"meaning" "abandoning evil" and "meditation." It comes from Ancient Indian philosophy. This method of "sitting in silence with Pranayama and meditating to practice nyujo (Buddhist Sanskrit, Chinese: Ruding, 入定)" was introduced into Buddhism as a practice to suppress one's desires, introspect upon one's actions and problems, and keep one's heart from any evil from in the outside world. This practice is called "sitting meditation" or "meditation" in Buddhism, and through introspection (cognitively) and meditation (behaviorally), one can free oneself from external influences and make the mind peaceful. Rinzai (Lin-ji) Zen is a widely known and important school of Buddhism that advocates the practice of meditation. In modern society, there are many kinds of meditation. Some of them are related to Zen Buddhism. Others seem to be unrelated to it. This article explores Zen and the psychological significance of meditation as related to believing. It focuses on the meditation related to Zen Buddhism. However, there are other forms of meditation as well, because not all practitioners are Buddhists, although they must have beliefs about meditation for it to be effective and significant for them.

Subsections relevant to the subject: Zen and meditation

The origin and spread of Zen

Zen Buddhism, which advocates the practice of Zen meditation, is an essential school of Buddhism. Zen Buddhism was founded by the first Bodhidharma, who came to China from India, to the sixth master Huineng 慧能, who developed five schools and seven sects. It became the primary school of Chinese Buddhism after the Mid and Late Tang Dynasty. But after the Southern Song Dynasty, only two schools (Rinzai and Caodong 曹洞) were prevalent and brought to Japan, while the rest were not transmitted to subsequent generations. As an important school of Japanese Zen, Rinzai Zen originated mainly from the Lin-ji branch of Chinese Zen Buddhism (Chinese: Lin-ji zong, 临济宗). This meant that a disciple's satori could be directly imparted and received through the teacher's word. Some religious scholars reiterate the well-worn truism that the scholar

has no direct access to the inner experience of the mystic but only access to a verbal or written account of this experience. But this claim relies on a linguistic assumption that may not be shared by the mystics themselves (Johnson, 2017).

The history of Zen meditation went from Indian Zen to Chinese Zen, and then from Chinese Zen to Japanese Zen, which then evolved into a step-by-step development of Zen from the West to the East. Just as there is no such thing as "religion" in the singular, but there are many things called "religions" based on the myriad variations in which beliefs are acquired and evolve and incorporated into one's life, so also there is no "meditation" in the singular but there are a variety of "meditations," both religious and non-religious, that both manifest and modify beliefs. They exist because they serve the needs of human spiritual and mental health. In the process of Zen's Westward journey, the renowned scholar Suzuki Daizo played a great role in the introduction and promotion of Zen Buddhism in Europe and America. He once said, "Zen is essentially the art of gaining insight into the nature of human life, which points the way from slavery to freedom" (Erich et al., 1987). Some meditation practices are related to Buddhism and some are not. This distinction will be discussed next.

The psychological significance of meditation related to Buddhist beliefs

During modern times, "sitting meditation" became a cultural symbol of traditional eastern psychotherapy after Zen Buddhism was introduced to the Western world by Japanese scholars. Meanwhile, in the Asian regions, such as in some cities of China, where psychology is in the process of being established as a new and developing discipline, there are a number of psychologists who draw on traditional Buddhist theories to apply them to their clinical practice in order to promote mental health. Because of the reciprocal benefit between various forms of meditation and the discipline of psychology in recent decades, the ancient Buddhism has had to adapt itself to meet the needs of modern people for spiritual health in the face of transformations in modern society. Modern Buddhist priests explain and practice Buddhist meditation from the perspective of psychology and mental health believing that it helps people relieve the pain of psychological and mental illnesses. There are some courses on the psychological bases of meditation for monks or nuns in some Buddhist colleges and institutes in Taiwan, so that they could provide psychological counseling services for their society after training. Based on their own experience in meditation, some senior monks have been able to help people become happier physically and mentally and to rid themselves of worries and fatigue by sitting in meditation and teaching sutras to the faithful (Chen and Deng, 2001).

Sitting meditation requires both "inward introspection" (which includes one's self-examination of not only one's own beliefs, but also of one's own believing process), freed from outside interference to achieve enlightenment, and adjustment of the body and mind through "meditation" to achieve a healthy state. The practice was developed in Japan into "Inner Vision Therapy," which was absorbed by Europe more than 1,000 years later. In Germany, Hewels created the "Self-Therapy"—a psychotherapy based on sitting in meditation that had an influence on psychological and other helping techniques around the world. Coleman suggests that Buddhism's doctrine of no-self can help solve the fragmentation of self that can occur in postmodern societies (Coleman, 2001). Buddhist meditation practices underpin some cognitive behavioral therapies (which involve working with a client to change his or her negative beliefs about self) currently used in psychology to treat mental disorders (Carmody et al., 2008).

The Swiss psychoanalyst Jung concluded that "sitting meditation" has an important role in psychology. Since he wrote, some psychologists studied "meditation" and summarized the five techniques of "sitting in meditation," which are equivalent to the "five dharma-paryaya" (dharma-gate, Chinese: Famen, 法门) in Buddhism: (a) Pranayama (to regulate the breath, Chinese: Tiaoxi, 调息); (b) Asubha (to reflections on repulsiveness, Chinese: Bujing, 不净); (c) Karunā (compassion or mercy, Chinese: Cibei, 慈悲); (d) Hetu-pratyaya (Karma, Chinese: Yinyuan, 因缘). For all the above, the practitioner must use reason and calmness to understand the cause and effect of things and use reason to overcome evil thoughts. Finally, (e) Patha (chanting, Chinese: Songnian, 诵念), which means imagining the Buddha's sitting posture and chanting his name to get rid of fear, conquer desires, and reach a state of purity and peace. This method is also used in suggestion therapy, such as with self-referral words, imagination, etc., (Xu, 2007). Success in all five of the above cases hinges on believing being central to the process of meditation; and in this way it demonstrates its significance.

The psychological significance of meditation in non-religious believing

In recent years, sitting meditation has become a fashion in some regions, but not necessarily practiced as part of one's religiousness. The psychological and mental health fields have introduced parts of oriental Zen theory into their own disciplinary areas and developed a body of knowledge and discourse that has been de-contextualized over a long period of time so that it longer puts an emphasis on having originated from Rinzai Zen Buddhism. For example, Mindfulness meditation, in particular, has often been as part of psychotherapy with little or no attention to its religious roots or underpinnings. Current teaching and practice of

mindfulness are limited to focus on one's personal wellbeing, which has therapeutic benefits but is only a small part of the wisdom, philosophy, and fundamental principles of Buddhist teaching and practice. The use of mindfulness in such a limited fashion has caused concern and raised questions. Nonetheless, mindfulness is more than that which meets the breath. It has a more powerful message and a deeper purpose; it is fundamentally about the human condition and liberation (Manikam, 2016).

Modern people are more likely to use sitting meditation as an exercise related to the promotion of spiritual and mental health. Shanghai Yufo Temple 上海玉佛寺in China has held a free 2-day sitting meditation activity for many years, and participants do not need to have Buddhist beliefs, nor do they need to pay any fees, which are open to the public and very popular among urban young people (Du, 2015). Further research is needed to understand the implications of isolating such practices from their historical roots and religious contexts.

Many modern psychologists have highly evaluated the psychological significance of "sitting meditation" and generally believe that "sitting meditation" is not only a mystical religious system but also a psychological practice in the modern sense. Van Gordon et al. (2017) mentioned that a number of studies had investigated the utility of a secular (but Buddhist based) 8-week intervention known as meditation awareness training (MAT), which assigns to training participants in the concept and practice of emptiness (as well as other Buddhist meditative and spiritual techniques). Findings—including from clinical case studies as well as randomized and non-randomized controlled trials—have shown that MAT can improve: (a) work-related stress, (b) stress, anxiety, and depression, (c) workaholism, (d) co-occurring schizophrenia and pathological gambling, and (e) job satisfaction, organizational citizenship, and job performance (Shonin and Van Gordon, 2015).

By what processes does meditation have these effects? Such effects are related to what someone comes to believe and how they come to believe it. Meditation could also be related to the processes through which someone's beliefs change over time—for example, from pre-to-post meditation. It is exactly such hoped-for changes that many people practice mindfulness meditation. And if the changes that they want happen, they come to "believe in" meditation more than they did before—which illustrates how the believing process is not static, but fluid. This illustrates a different kind of significance of mediation, i.e., that practicing meditation involves processes of believing at its core.

Other significance of meditation related to the believing

Scholars have increasingly become interested in silence and its role in social interaction during Buddhist meditation (Kurzon, 2007, 2011; Ephratt, 2008). Fennell (2012) analyzed the silences created in an unprogrammed Quaker Meeting for Worship as well as the meditation practices of three Buddhist groups: a sitting group associated with a Vipassana organization, a Zen sitting group under the leadership of a person ordained by Thich Nhat Hanh, and a temple under the leadership of a Zen priest of the Soto school. Such participants were able to get something out of sitting with the groups because the leaders treated learning as a shared experience. They were willing to listen to and learn with other participants, and created spaces for participants to share their perspectives. Some researchers consider silence in a spiritual context as fostering intrapersonal communication (Smith et al., 2010). Such silence operates as a conduit for people to communicate with God, deities, nature, or the "self" (Bauman, 1993; Jaworski, 1993; Dandelion, 1996; Lightstone et al., 2006; Levine, 2008), but not necessarily other people. Quakers feel they can communicate with God (Fox, 1671), and Buddhism "has in some cases been reconfigured as a technique for self-discovery" (McMahan, 2008). Even when silence is seen as active, it may be construed as functioning to turn people inward and away from each other. Instead, meditation demonstrates how silence for these groups is not necessarily socially isolating. For instance, participants not only learned, practiced, and reflected on silence together, but drew connections between the practice of silence and changes in everyday social interactions, including bringing compassion into them (Fennell, 2012).

Discussion and conclusion

Meditation has its roots in Zen Buddhism, but later developed in a number of different variations. This article explores "sitting meditation," not only as a mystical religious practice but also as a spiritual practice related to the believing processes understood in modern psychology. Buddhism is full of educational psychology theories, and the Buddha himself was an effective educator. The Buddha made meticulous observations of the human mind, revealed the incredible potential of the human heart, and proposed a series of operational techniques for self-knowledge, self-purification of the mind, and development of the potential of the mind (Chen, 2006). Besides those related to Buddhism, there are other types of meditation whose practitioners engage in believing in order for meditation to be effective and meaningful to them. They may not be Buddhist believers, but believing certain things is an intimate component in the very process of performing the practice of meditation, for the practice can be significant for them only if it is underpinned by believing. The dialogue between believing and psychology makes it possible to tackle more issues to be addressed in the fields of religious spirituality and mental health.

Author contributions

The author confirms being the sole contributor of this work and has approved it for publication.

Funding

This article was supported by the National Social Science Fund of China under the project "A study of the transmission patterns and development paths of Chinese culture among new Chinese immigrants in Canada" (Project No. 22BMZ141) and the National Social Science Fund of China under the major project "Research on the discourse system of religious sociology with Chinese characteristics and its local knowledge structure" (Project No. 18ZDA230).

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Bauman, R. (1993). "Speaking in the light: the role of the quaker minister," in *Explorations in the Ethnography of Speaking*, eds R. Bauman and J. Sherzer (New York, NY: Cambridge University Press), 144–160. doi: 10.1017/CBO9780511611810.012

Carmody, J., Reed, G., Kristeller, J., and Merriam, P. (2008). Mindfulness, spirituality, and health-related symptoms. *J. Psychosom. Res.* 64, 393–403. doi: 10.1016/j.jpsychores.2007.06.015

Chen, B. 陈兵(2006). *The Wisdom of Buddha* 佛陀的智慧. Shanghai: Shanghai Guji Press, 294.

Chen, B., and Deng, Z. 陈兵、邓子美 (2001). *Chinese Buddhism in the Twentieth Century* 二十世纪的中国佛教. Beijing: Nation Press, 130.

Coleman, J. W. (2001). *The New Buddhism: The Western Transformation of an Ancient Tradition*. New York, NY: Oxford University Press.

Dandelion, P. (1996). *A Sociological Analysis of the Theology of Quakers: The Silent Revolution*. Lewiston: Edwin Mellen Press.

Du, L.-J. 杜立婕(2015). A sociological interpretation of the meaning of Buddhism in "urbanization 佛教之于"城市化"意义的社会学诠释". *Urban Cult. Res.* 都市文化研究 11, 205–214.

Ephratt, M. (2008). The functions of silence. *J. Pragm.* 40, 1909–1938. doi: 10.1016/j.pragma.2008.03.009

Erich, F., Richard, M., and and, D. T., Suzuki 铃木大拙(1987). *Zen Buddhism and Psychoanalysis.* Transl. by L. Wang and C. Feng. Guizhou Guizhou People's Publishing House, 155–164

Fennell, D. (2012). Explorations of silence in the religious rituals of Buddhists and Quakers. *Religion* 42, 549−574. doi: 10.1080/0048721X.2012.704157

Fox, G. (1671). *Some Principles of the Elect People of God in Scorn Called Quakers.* Available online at: http://dqc.esr.earlham.edu:8080/xmlmm/docButton?XMLMMWhat=tocandXMLMMWhere=1&XMLMMBeanName=toc1&XMLMMNextPage=/printBuiltTocPage.jsp&XMLMMCheck=6.1.1.485.477

Jaworski, A. (1993). *The Power of Silence*. Newbury Park, CA: Sage.

Johnson, C. D. (2017). *Zen and the Unspeakable God.* University Park, PA: Pennsylvania State University Press, 522–525.

Kurzon, D. (2007). Towards a typology of silence. *J. Pragm.* 39, 1673–1688. doi: 10.1016/j.pragma.2007.07.003

Kurzon, D. (2011). On Silence. *J. Pragm.* 43, 2275–2277. doi: 10.1016/j.pragma.2010.11.011

Levine, M. (2008). Prayer as coping: a psychological analysis. *J. Health Care Chaplain.* 15, 80–98. doi: 10.1080/08854720903113424

Lightstone, J. N., Bird, F. B., and Fishbane, L. S. (2006). *Ritual and Ethnic Identity: A Comparative Study of the Social Meaning of Liturgical Ritual in Synagogues.* Waterloo, ON: Wilfrid Laurier University Press, 23–52.

Manikam, R. (2016). Buddhist foundations of mindfulness. *Mindfulness* 7, 1005–1008. doi: 10.1007/s12671-016-0523-y

McMahan, D. L. (2008). *The Making of Buddhist Modernism*. New York, NY: Oxford University Press, 184. doi: 10.1093/acprof:oso/9780195183276.001.0001

Shonin, E., and Van Gordon, W. (2015). Managers' experiences of meditation awareness training. *Mindfulness* 6, 899–909. doi: 10.1007/s12671-014-0334-y

Smith, P. H., Murray, C. E., and Coker, A. L. (2010). The coping window: a contextual understanding of the methods women use to cope with battering. *Violence Victims* 25, 18–28. doi: 10.1891/0886-6708.25.1.18

Van Gordon, W., Shonin, E., and Griffiths, M. D. (2017). Buddhist emptiness theory: implications for psychology. *Psychol. Relig. Spiritual.* 9:309. doi: 10.1037/rel0000079

Xu, G. 徐光兴(2007). *Mental Zen: Oriental Psychological Therapy* 心理禅—东方人的心理疗法, 1st Edn. Shanghai: East China Normal University Press; Wenhui Publishing House.

frontiers | Frontiers in Psychology

TYPE Opinion
PUBLISHED 04 October 2022
DOI 10.3389/fpsyg.2022.971284

Check for updates

OPEN ACCESS

EDITED BY
Rüdiger J. Seitz,
Heinrich Heine University of Düsseldorf, Germany

REVIEWED BY
Nicholas John Shaman,
University of Houston–Clear Lake, United States

*CORRESPONDENCE
Courtney Applewhite
applewhite@ucsb.edu

SPECIALTY SECTION
This article was submitted to Theoretical and Philosophical Psychology, a section of the journal Frontiers in Psychology

RECEIVED 17 June 2022
ACCEPTED 06 September 2022
PUBLISHED 04 October 2022

CITATION
Applewhite C (2022) Believing and the disposal of bodies after death. *Front. Psychol.* 13:971284. doi: 10.3389/fpsyg.2022.971284

COPYRIGHT
© 2022 Applewhite. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Believing and the disposal of bodies after death

Courtney Applewhite*

University of California, Santa Barbara, Santa Barbara, CA, United States

KEYWORDS

death and dying, meaning making process, death, global meaning, non-religious people, secular people, ritual, promiscuous teleology

Introduction

One of the first realities of death is the presence of the physical dead body. In studies of death in fields such as neuroscience or psychology, we focus on the biological basis or clinical implications of grief and bereavement (Neimeyer, 2004). But grief and bereavement are a response to death and, thus, a dead body. Psychologists have paid little attention to the material body of the deceased and the meaning-making processes associated with it (White et al., 2016). I will approach the meaning making process from two perspectives: the first, how our beliefs affect the appraisal of the deceased body of another person, perhaps a loved one, and the second, how those same beliefs may be similar or different in how we appraise our own deceased body as we consider what we want done with it after our own death.

I aim to shift the focus from the experience of grief specifically to the construction of meaning in relation to the physical body after death or, as they say in the profession, the choice of body disposition. To do so, we can turn to Park (2010) integrated notion for meaning making as it relates to how people might appraise a dead body. While cremation and conventional burial practices (in which a person is embalmed, placed in a casket, and then buried in a vaulted grave) are still the most popular choices for bodily disposition in the United States, other practices like green or natural burial, alkaline hydrolysis, and natural organic reduction are becoming more widely available and requested, particularly by non-religious people. I argue that in choosing these alternative methods of bodily disposition for themselves or for their loved ones, non-religious people are enacting a different kind of belief by simultaneously recognizing the materiality of the body and ascribing value and meaning to it from spiritual, environmental, and/or cultural perspectives. This way of viewing appraisal will draw from a relational-deictic framework and consider how people often hold simultaneous and sometimes contradictory appraisals.

Integrated meaning making and global meaning

Our global meaning systems are made up of beliefs, goals, and subjective feelings as well as internal representations of desired processes, events, or outcomes (Park, 2010). These beliefs, which take many forms, e.g., religious, spiritual, ethical, or material, guide how we make sense of the world around us. As we encounter stressful moments throughout life, we must make sense of them in some way; we must ascribe meaning

to them. In forming meaning in these moments, called situational meaning, we must attempt to either assimilate–change our view of the situation–or accommodate–change our global meaning. This process is not simply cognitive, but also relies on emotional processing. For many, viewing a dead body and/or making decisions about how to dispose of it is one of those stressful moments.

The meaning-making process begins with the physical reality of the deceased body. There are two separate ways to approach the idea of the deceased body. One is through the experience people have in choosing a disposition method for another person at the time of their death. Another is making a future appraisal of your own dead body. Keeping these two kinds of events in mind, we can consider how non-religious individuals in the United States react to death when they often lack the institutionalized rituals and routines for responding to a death that religious institutions typically provide (MacMurray and Fazzino, 2017). For example, if you are Catholic in the United States, most of the time from the guidance of your parish priest you would choose to bury a loved one or to personally be buried in a Catholic cemetery. There are rituals and masses associated with death that are congruent with your global meaning structure. The growing number of people in the United States that do not subscribe to any religious tradition do not have the rituals, rites, and ceremonies around bodily disposition and therefore do not necessarily move as smoothly through the meaning making process when deciding what to do with the deceased (Kosmin et al., 2009; MacMurray and Fazzino, 2017; Smith and Cragun, 2019). As such, they must find new ways to make meaning that can help them make sense of death and dispose of the body in a way that makes sense to them *as secular people*.

Shifting disposition methods in the United States

The self-understanding that people brought to the dead used to be relatively consistent. In the antebellum United States, funeral services were presided over by Christian clergy, occurred soon after death, and were rarely attended by anyone other than immediate family (Laderman, 1996). The family's concerns centered on religious beliefs, namely the deceased's soul and whether it was bound for heaven. Both the form and content of these services shifted after the Civil War. After the 1860s, technological advancement allowed for the professionalization of the funeral industry, notions about the ontology of heaven and hell changed, and there was extended time before burial (Laderman, 1996; Prothero, 2001). Together, this meant that the presentation of an embalmed corpse in an open casket in the context of a religious service became popular in the United States and, until about the 1960s, was the norm.

Cremation increased in popularity in the 1960s in response to Jessica Mitford's aggressive critique around pricing in the funeral industry, the Catholic Church lifting the cremation ban, and the rise of the counterculture (Prothero, 2001). But cremation numbers remained relatively low even in the 1990s. Since 2005, the Cremation Association of North America has been collecting data on cremation rates based on their members and affiliates. Their data demonstrate that cremations have been climbing steadily only to surpass the number of burials in 2015 and come to represent more than 50% of dispositions in 2016. Cremation rates have been increasing steadily over the past several decades; there has not been an overwhelming jump or shift since that period in the 1960s (Kemmis, 2021). This trend is like much of Europe, although the numbers in Europe have increased to even greater heights in many places. Cremation is not the only "new" practice growing in popularity. Other, more diverse options are also arising.

Non-religious people and disposition choice

Interest in alternative forms of disposition is on the rise (National Funeral Directors Association, 2022). The National Funeral Directors Association found that over 60% of people expressed interest in green burial in particular. This is also demonstrated by the increase in the availability of alternatives to cremation and conventional burial like alkaline hydrolysis, or water cremation, and natural organic reduction (i.e., human composting) as well as body donation, which has long been favored by the non-religious–specifically Atheists (Copeman and Quack, 2015). As the numbers of non-religious people rise in the United States, there may be greater opening to exploring utilitarian or other disposition options that do not have religious associations (Marsh, 2021). One of the values that may be applied to disposition is a concern for the environment (van Mulukom et al., 2022). Even though there is an important overlap in environmental protection impulses in both religious and non-religious individuals, in death these innovative and "green" practices are relatively free of religious baggage (Beaman, 2017).

I suggest that the focus on the environmental impact of their own deceased body indicates an appraisal of that body as less sacred and more material. It may also demonstrate how secular people are creating practices for themselves around death that reflect a change in belief. Non-religious people often do not believe that the body is critical to any kind of afterlife, so they are reframing the appraisal. But their intentional choices suggest that non-religious people are actively seeking actions associated with their understanding of death and the dead body that align with their values and beliefs in life, even those not religion related.

How belief influences disposition: Relational-deictic framework

Alongside a change in disposition preference, we are also seeing an increasing number of people who do not believe in a particular afterlife. One-in-six Americans do not believe in any afterlife at all (Pew, 2021). When discussing non-religious individuals, I am describing a subsection of this identification that have a materialist worldview and are part of this group that does not believe in a particular afterlife. This certainly does not describe the plurality of people who identify as non-religious. But within this group, we see this belief pattern that has an increasing number of people believing that death is the final end and wanting these environmentally friendly disposition practices.

To further consider how non-religious people might be thinking about the dead body, I propose that we turn to a "relational-deictic" interpretation of the physical world. Appraisal through a relational-deictic interpretation is an alternative to promiscuous teleology or the bias toward purpose-based reasoning (ojalehto et al., 2013). Teleological explanations become promiscuous when applied to natural objects rather than artifacts. For example, the clouds exist because they provide shade. These biases exist in humans from childhood into adulthood, although often not explicitly, and religion can be thought of as a product of this kind of reasoning (Kelemen et al., 2013). In this example, clouds would be a product of divine creation. As an alternative, the relational-deictic framework takes into account the importance of relational and ecological reasoning as well as the points of view within that relation (ojalehto et al., 2013). Rather than taking an intentional design stance with respect to natural forms, which presupposes that the purpose arises from a designer or a sole source, the relational-deictic stance assumes that purposes come from multiple sources and therefore purpose arises from the perceiver's *sense* of purpose. This becomes critical when considering the reality of the dead body.

While ojalehto et al. (2013) point out that it is Indigenous populations in the United States that most clearly display this kind of cognitive approach, I argue that this reasoning can be a better way to talk about the cognition of the non-religious people who are choosing alternative disposition methods. The key relationship is between the living and the dead body. From this relationship emerges additional connections because both are part of the natural world, but one will continue on existing in that natural world (the living) and another will decompose. So, although the living recognizes the dead body as part of that cycle of nature, as material, it remains important because of the ongoing connection with the person the body used to be. From a teleological perspective, death may be "part of God's plan," or it may "happen for a reason," but these phrases are not only laden with religious connotation, but also fundamentally incongruent with many non-religious people who subscribe to materialist worldviews. And yet, even if you do not believe in an afterlife, or believe in a soul, or the sacrality of the body from a religious or spiritual perspective, which many non-religious people do not, you still maintain a relationship to that person who has died. When we appraise the deceased body of another person, particularly a loved one, we see that body as more than simply material.

Using a relational-deictic framework would shift the language associated with death, particularly when choosing these alternative forms of disposition, to phrases like, "the body is part of a natural cycle," or "the person's body will go back to the earth to support it." Hence, the body becomes a key part of a purpose or cycle without the baggage of promiscuous teleology. This framework is critical to the non-religious because they lack theistic global meaning about the purpose of the dead body, or death in general, and the traditional forms of disposition do not necessarily allow the same kind of reasoning. Conventional burial and cremation do not lend themselves to this natural return to the earth as easily.

Discussion

Relational frameworks use sophisticated ecological reasoning that is particular to Indigenous communities, but this argument suggests a biological basis to this process, which may be why we apply it in considering the cognitive underpinnings of non-religious people. And a possible pathway for this cognitive logic is that there is a coexistence of natural and supernatural explanations for things within people's minds for both themselves and for others (Legare and Shtulman, 2018). The physical body can represent the vessel for the soul, but it is also a biological fact. The physical body may both be important and not be important. Relational frameworks are most important for the bodies of those to whom people are related. This explains why some people tend to be quite flippant about their own death and their own bodies after death but would very rarely be as cavalier about a loved one's body. This suggests that global meaning is not necessarily a fixed or total system, but it can have contradictory pieces that are held simultaneously and constantly shifting.

Global meaning systems can consider several different causalities and with several contradictory points of view within the same person, but then we return to the physical reality of the dead body. This short article suggests that as non-religious people are moving away from the dominant religious narratives that provide meaning and structure around the dead body for both themselves and others, they are introducing other kinds of meaning. These meanings include values and beliefs around environmentalism, secularism, economics, or tradition outside of religion, which has perhaps influenced the growing numbers

of people who are interested in green burial, natural organic reduction, and other means of bodily disposition. For them, the body is not sacred in the religious sense, but is indeed value-laden from a relational and natural perspective. Many, despite their non-belief, still ascribe a specialness to the deceased body, a cognitive and emotional response that bears further investigation. Future research may directly investigate how non-religious people think about the deceased body and relate to other phenomena around death, not limited to after-death communications, or sensing presences, and the experiences of grief.

Author contributions

The author confirms being the sole contributor of this work and has approved it for publication.

Funding

This paper was funded by Dr. Rüdiger Seitz, *via* the Volkswagen Foundation, Siemens Healthineers, and the Betz Foundation.

Acknowledgments

The author thanks Ann Taves and Ray Paloutzian for their review and comments on the early drafts and revision of this article. Thank you to the several colleagues who offered comments at the Credition conference, which has shaped my update of this article from our gathering.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Beaman, L. G. (2017). Living well together in a (non)religious future: contributions from the sociology of religion. *Sociol. Relig.* 78, 9–32. doi: 10.1093/socrel/srw054

Copeman, J., and Quack, J. (2015). Godless people and dead bodies: materiality and the morality of atheist materialism. *Soc. Anal.* 59, 40–61. doi: 10.3167/sa.2015.590203

Kelemen, D., Rottman, J., and Seston, R. (2013). Professional physical scientists display tenacious teleological tendencies: purpose-based reasoning as a cognitive default. *J. Exp. Psychol. Gen.* 142, 1074–1083. doi: 10.1037/a0030399

Kemmis, B. (2021). *The Cremation Landscape. Cremation Association of North America (CANA)*. Available online at: https://www.cremationassociation.org/blogpost/776820/370274/The-Cremation-Landscape (accessed May 11, 2021).

Kosmin, B., Keysar, A., Cragun, R., and Navarro-Rivera, J. (2009). *American Nones: The Profile of the No Religion Population [A Report Based on the American Religious Identification Survey 2008]*. Trinity College.

Laderman, G. (1996). *The Sacred Remains: American Attitudes Toward Death, 1799-1883*. New Haven, CT: Yale University Press.

Legare, C. H., and Shtulman, A. (2018). "Explanatory pluralism across cultures and development," in *Metacognitive Diversity: An Interdisciplinary Approach* (Vol. 1), eds J. Proust and M. Fortier (Oxford, UK: Oxford University Press). doi: 10.1093/oso/9780198789710.003.0019

MacMurray, N. J., and Fazzino, L. L. (2017). "Doing death without deity: constructing nonreligious tools at the end of life," in *Organized Secularism in the United States, eds* R. T. Cragun, C. Manning, and L. L. Fazzino (Berlin: De Gruyter), 279–300. doi: 10.1515/9783110458657-014

Marsh, H. M. (2021). The green new death: a legislative framework to promote and legalize green funerary alternatives notes. *George Washington J. Energy Env. Law* 12, 124–136.

National Funeral Directors Association. (2022). *Statistics*. National Funeral Directors Association. Available online at: https://nfda.org/news/statistics (accessed April 15, 2022).

Neimeyer, R. A. (2004). Report on bereavement and grief research. *Death Stud.* 28, 491–575. doi: 10.1080/07481180490461179

ojalehto, B., Waxman, S. R., and Medin, D. L. (2013). Teleological reasoning about nature: Intentional design or relational perspectives? *Trends Cogn. Sci.* 17, 166–171. doi: 10.1016/j.tics.2013.02.006

Park, C. (2010). Making sense of the meaning literature: an integrative review of meaning making and its effects on adjustment to life events. *Psychol. Bull.* 136, 257–301. doi: 10.1037/a0018301

Pew. (2021). *How COVID-19 Has Strengthened Religious Faith. Pew Research Center's Religion and Public Life Project*. Available online at: https://www.pewforum.org/2021/01/27/more-americans-than-people-in-other-advanced-economies-say-covid-19-has-strengthened-religious-faith/ (accessed January 27, 2021).

Prothero, S. (2001). *Purified by Fire: A History of Cremation in America*. Oakland, CA: University of California Press. doi: 10.1525/9780520929746

Smith, J. M., and Cragun, R. T. (2019). Mapping Religion's Other: A Review of the Study of Nonreligion and Secularity. *J. Sci. Study Relig.* 58, 319–335. doi: 10.1111/jssr.12597

van Mulukom, V., Turpin, H., Haimila, R., Purzycki, B. G., Bendixen, T., Kundtov, E., et al. (2022). What do nonreligious nonbelievers believe in? secular worldviews around the world. *Psychol. Relig. Spiritual.* 15, 480. doi: 10.1037/rel0000480

White, C., Fessler, D. M. T., and Gomez, P. S. (2016). The effects of corpse viewing and corpse condition on vigilance for deceased loved ones. *Evolut. Hum. Behav.* 37, 517–522. doi: 10.1016/j.evolhumbehav.2016.05.006

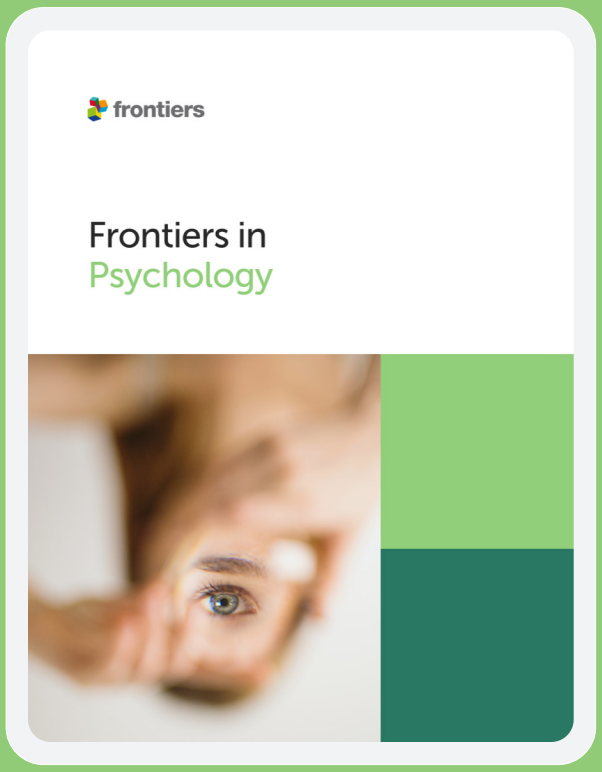
Frontiers in Psychology

Paving the way for a greater understanding of human behavior

The most cited journal in its field, exploring psychological sciences - from clinical research to cognitive science, from imaging studies to human factors, and from animal cognition to social psychology.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact