Research Topics

How emotion relates to language, memory, and cognition

Edited by
Michael J. Cortese, Maya Michelle Khanna and Francesca M. M. Citronv

Published in
Frontiers in Communication
Frontiers in Psychology

ISSN 1664-8714
ISBN 978-2-83252-023-9
DOI 10.3389/978-2-83252-023-9

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

How emotion relates to language, memory, and cognition

Topic editors
Michael J. Cortese — University of Nebraska Omaha, United States
Maya Michelle Khanna — Creighton University, United States
Francesca M. M. Citron — Lancaster University, United Kingdom

Citation
Cortese, M. J., Khanna, M. M., Citron, F. M. M., eds. (2023). *How emotion relates to language, memory, and cognition*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-83252-023-9

Table of contents

frontiers | Frontiers in Communication

TYPE Editorial
PUBLISHED 13 March 2023
DOI 10.3389/fcomm.2023.1170912

OPEN ACCESS

EDITED AND REVIEWED BY
Xiaolin Zhou,
Peking University, China

*CORRESPONDENCE
Francesca M. M. Citron
✉ fmm.citron@gmail.com

SPECIALTY SECTION
This article was submitted to Language Sciences, a section of the journal Frontiers in Communication

RECEIVED 21 February 2023
ACCEPTED 28 February 2023
PUBLISHED 13 March 2023

CITATION
Citron FMM, Cortese MJ and Khanna MM (2023) Editorial: How emotion relates to language, memory, and cognition. *Front. Commun.* 8:1170912. doi: 10.3389/fcomm.2023.1170912

Editorial: How emotion relates to language, memory, and cognition

Francesca M. M. Citron[1]*, Michael J. Cortese[2] and Maya M. Khanna[3]

[1]Department of Psychology, Lancaster University, Lancaster, United Kingdom, [2]Department of Psychology, University of Nebraska Omaha, Omaha, NE, United States, [3]College of Arts and Sciences, Psychology, Creighton University, Omaha, NE, United States

KEYWORDS

emotive words, affective priming, sentiment analysis, mental health, discourse, pragmatics, emotion

Editorial on the Research Topic
How emotion relates to language, memory, and cognition

In cognitive neuroscience, the role of the cerebral cortex in sensation and movement was known long before the emotion neural network (e.g., Libet et al., 1964). Similarly, theoretical models of reading and empirical research on word recognition had not considered the relevance of words' emotive content until about 20 years ago. Then, research supporting the automatic vigilance hypothesis (Pratto and John, 1991) and subsequent megastudies showed that affective dimensions such as emotional valence and arousal affect word processing, above and beyond sub-lexical, lexical and semantic word properties (e.g., Larsen et al., 2006; Kuperman et al., 2014; Cortese and Khanna, 2022). Using a similar approach, Grzybowski et al. generated a database of state, trait, and hybrid positive and negative Polish adjectives of moderate to high arousal, which complements existing Polish word databases, and can be used alongside them to create personality and mood questionnaires as well as for other affective language research, more generally.

Concurrently, neurophysiological research showed that emotive words can trigger fast and automatic activation of the emotion neural network, similarly to actual threatening objects or distressful scenes. For example, evolutionary-relevant words are distinguished from neutral words around 200–300 ms, when we match word forms to our mental lexicon, and long before we gain full access to a word's meaning (around 400 ms; e.g., Kissler et al., 2007; Citron, 2012).

Considering these findings, single words represent excellent tools to investigate and devise treatments for different psychopathologies. For instance, in the emotional Stroop task, words can activate disorder-related concepts (Khanna et al., 2016). Furthermore, texts or discourse describing personal situations can reveal information on mental health conditions (e.g., Herbert et al., 2019). In particular, through sentiment analysis, the use and frequency of certain linguistic features can indicate different psychopathologies. Interestingly, Du could predict mental states of writer Virginia Wolff from her diary and biography. Sentiment analysis represents a powerful tool not only with regard to literary texts, spontaneous speech and mental health; it can also be used to analyze political speech. In fact, Whissell identified changes and stable features over time in US presidential candidates' nomination acceptance speeches. Further to literary texts, Hugentobler and Lüdtke investigated the effect of semantic cohesion on aesthetic appreciation of poems. Aiming to isolate semantic cohesion from other potential sources of appreciation, they presented word lists as modern micro poems to participants, who could more easily understand, appreciate and extract underlying concepts from cohesive micro-poems, as evidenced by several explicit and implicit measures.

Other empirical contributions in this Research Topic were concerned with more theoretical or conceptual distinctions with regards to affective language. In a novel ERP contribution, Wu et al. expanded their previous work on affective priming to explore the conceptual distinction between emotion-label and emotion-laden words originally posited by Altarriba and Basnight-Brown (2011); they tested the hypothesis that emotion-laden words cannot prime emotion-label words and provided a detailed timeline of these priming effects through ERPs. A very different contribution, also based on priming, comes from Rohr and Wentura, who reviewed and critically evaluated 20 years of research on evaluative priming paradigms, to then present a new model based on short-term memory representations. One of their conclusions was that the prime's emotional valence is automatically activated only if relevant to task goals.

However, language is first and foremost humans' means of communication, and verbal interactions are based on the comprehension of meaning in context. Three contributions go beyond single-word processing to examine the relationship of emotion and language in discourse. Struiksma et al. explored the electrophysiological responses to repeated insults toward the participant or a third party, comparing them with compliments and neutral statements, and revealing a very interesting picture. Israel et al. measured eye-movements and pupil dilation during the comprehension of humorous discourse, providing a timeline of situation model revision; the study also revealed an additional affective reaction compared to non-humorous discourse. Finally, Lai et al. reported an ERP study showing participants' mood or affective state can influence conversational exchanges. In particular, mood had differential effects on discourse comprehension at late processing stages, when either world knowledge, discourse context or both are used to make sense of the discourse; mood did not affect meaning retrieval *per se*.

Author contributions

FMMC drafted the editorial. MMK and MJC further contributed to it. 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

Altarriba, J., and Basnight-Brown, D. M. (2011). The representation of emotion vs. emotion-laden words in English and Spanish in the affective simon task. *Int. J. Bilingual.* 15, 310–328. doi: 10.1177/1367006910379261

Citron, F. M. M. (2012). Neural correlates of written emotion word processing: A review of recent electrophysiological and hemodynamic neuroimaging studies. *Brain Lang.* 122, 211–226. doi: 10.1016/j.bandl.2011.12.007

Cortese, M. J., and Khanna, M. M. (2022). Relating emotional variables to recognition memory performance: a large-scale re-analysis of megastudy data. *Memory* 30, 915–922. doi: 10.1080/09658211.2022.2055080

Herbert, C., Bendig, E., and Rojas, R. (2019). My sadness - Our happiness: Writing about positive, negative, and neutral autobiographical events reveals linguistic markers of self-positivity and individual well-being. *Front. Psychol.* 9, 2522. doi: 10.3389/fpsyg.2018.02522

Khanna, M. M., Badura-Brack, A. S., McDermott, T. J., Shepherd, A., Heinrichs-Graham, E., Pine, D. S., et al. (2016). Attention training normalises combat-related post-traumatic stress disorder effects on emotional Stroop performance using lexically matched word lists. *Cogn. Emot.* 30, 1521–1528. doi: 10.1080/02699931.2015.1076769

Kissler, J., Herbert, C., Peyk, P., and Junghofer, M. (2007). Buzzwords. Early cortical responses to emotional words during reading. *Psychol. Sci.* 18, 475–480. doi: 10.1111/j.1467-9280.2007.01924.x

Kuperman, V., Estes, Z., Brysbaert, M., and Warriner, A. B. (2014). Emotion and language: Valence and arousal affect word recognition. *J. Exper. Psychol.* 143, 1065–1081. doi: 10.1037/a0035669

Larsen, R. J., Mercer, K. A., and Balota, D. A. (2006). Lexical characteristics of words used in emotional Stroop experiments. *Emotion* 6, 62–72. doi: 10.1037/1528-3542.6.1.62

Libet, B., Alberts, W. W., Wright, E. W. Jr., Delattre, L. D., Levine, G., and Feinstein, B. (1964). Production of threshold levels of conscious sensation by electrical stimulation of human somatosensory cortex. *J. Neurophysiol.* 27, 546–578. doi: 10.1152/jn.1964.27.4.546

Pratto, F., and John, O. P. (1991). Automatic vigilance: the attention-grabbing power of negative social information. *J. Person. Soc. Psychol.* 61, 380–391. doi: 10.1037/0022-3514.61.3.380

frontiers in Psychology

METHODS
published: 08 September 2021
doi: 10.3389/fpsyg.2021.683012

Check for updates

The Words of Affectivity. Affect, Category, and Social Evaluation Norms for 400 Polish Adjectives

Szczepan J. Grzybowski[1]*, *Miroslaw Wyczesany*[2], *Hanna Cichecka*[1] and *Aleksandra Tokarska*[1]

[1] *Institute of Applied Psychology, Faculty of Management and Social Communication, Jagiellonian University, Kraków, Poland,*
[2] *Institute of Psychology, Faculty of Philosophy, Jagiellonian University, Kraków, Poland*

Emotional adjectives can be grouped into two main categories: denoting and connoting stable (personality) traits and denoting and connoting transient (mood) states. They relate closely to the concept of affectivity, which is a pervasive tendency to experience moods of positive or negative valence. They constitute a rich study material for personality and affect psychology and neuroscience. Thus, this study was designed to establish a normed list of emotional adjectives with ratings encompassing four dimensions: emotional valence (positive or negative), emotional arousal (low-arousing or high-arousing), category (state, trait, and hybrid), and social judgment (competence, morality, and mixed). The adjectives were preselected based on previous broad Polish norming studies, personality and mood questionnaires, and a dictionary study. The results of the study were drawn from 195 participants who rated 400 adjectives that were chosen based on similar linguistic variables, such as frequency and word length. The dataset measures were proven to be stable and reliable. Correlations between the emotional valence and state-trait, valence and competence-morality, and emotional arousal and competence-morality dimensions were found. The study was successful in preparing a dataset of well-categorized (state, trait, and hybrid) positive and negative adjectives of moderate to high arousal ratings. Since the words were matched on linguistic variables, the set provided useful material that can be readily used for research into the effects of the category and emotional dimensions on language processing and as a basis for new personality questionnaires and mood checklists. The dataset could also be seen as a supplement for broader sets of published normed materials in Polish that link emotion and language.

Keywords: adjectives, affective norms, trait, state, word list, Polish language

OPEN ACCESS

Edited by:
Francesca M. M. Citron, Lancaster University, United Kingdom

Reviewed by:
Bogdan Wojciszke, University of Social Sciences and Humanities, Poland
Pilar Ferré Romeu, University of Rovira i Virgili, Spain

****Correspondence:***
Szczepan J. Grzybowski
sz.grzybowski@uj.edu.pl

Specialty section:
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

Received: *19 March 2021*
Accepted: *03 August 2021*
Published: *08 September 2021*

Citation:
Grzybowski SJ, Wyczesany M, Cichecka H and Tokarska A (2021) The Words of Affectivity. Affect, Category, and Social Evaluation Norms for 400 Polish Adjectives. Front. Psychol. 12:683012. doi: 10.3389/fpsyg.2021.683012

INTRODUCTION

The Concept of Affectivity

The ability to categorize and communicate the emotions of oneself and others is of paramount importance to human beings in both personal and social contexts. The categorization, experience, and communication of these emotions are mostly based on lexical means, i.e., words. According to the lexical hypothesis, which is well-established in personality psychology, the most relevant and important aspects of human experiences (traits and differences) become part of a language, with the aspects of utmost importance and usefulness forming single words (Cattell, 1943). A better understanding of the linguistic means and ways (personal, social, and culturally specific) related to emotional experience has been one of the goals of psychologists, linguists, and cultural anthropologists for decades (Wierzbicka, 1994, 1999). A psychological concept that can help us navigate through the existing theoretical and empirical richness and complexities is affectivity, which is a somewhat forgotten framework proposed by Tellegen (1985) and Tellegen and Waller (2008) and then elaborated on by Watson (Watson and Clark, 1984; Watson and Tellegen, 1985). Essentially, affectivity is a stable tendency to experience particular mood states. It can be seen as a personality trait that closely relates to and depends on mood. The affective state, i.e., mood, can be structured by itself in a two-dimensional, orthogonal fashion as a positive and negative affect (Watson and Tellegen, 1985).

By encompassing two mood-dispositional dimensions, affectivity is then evidenced by a pervasive predisposition to experience either negative mood states, termed as negative affectivity (NA), or positive ones, termed as positive affectivity (PA). It is worth noting that both of these dimensions influence the broader aspects of our lives, such as self-concept and cognition (Watson et al., 1988). A person characterized by high NA tends to often experience sadness, tension, anxiety, worry, hostility, and disgust, whereas a high PA person is more likely to experience feelings of joy, enthusiasm, interest, mental alertness, and energy regularly. Negative affectivity correlates positively with depression and anxiety, whereas PA correlates negatively with depression only (Watson et al., 1988). High PA interestingly reflects a sense of well-being and competence, which then forms the important aspects of self-perception, self-worth, and engagement with others (Wojciszke, 1997).

Affectivity as a personal trait is based on and conveyed by emotional experiences perceived in the self and others as regularly occurring states. The regularity takes the form of a stable tendency to experience a broad set of states with similar emotional attributes or attributes that are described on similar dimensions. Two of the most important and rudimentary of these attributes, i.e., dimensions, are emotional valence (or pleasure–displeasure) and arousal (or activation–deactivation). In fact, these two attributes combined form a core effect, a basic, pre-conceptual form of human experience (Russell, 2003, 2009). It is worth noting that affectivity, in a broad scope, can be seen as the source of the observed tendency and the result of it. In other words, e.g., a highly neurotic person will often and be more prone to experiencing feelings of worry, tension, and anxiety because of personal constituents. Conversely, a person that often happens to experience such states due to circumstances, work environment, and other situations can become neurotic or be described as neurotic by others. Regardless of the cause-and-effect description, affectivity points to the importance and intricacy of emotional experience related to both traits and states. How people categorize, encode, comprehend, and convey such experiences seem to be as worthwhile and engaging as an endeavor now as it was decades ago. The most common and substantial means of "dealing" (in the abovementioned terms) with such experiences are verbal stimuli of one particular grammatical class, namely, adjectives.

TABLE 1 | Descriptive statistics calculated for each dimension for all the intelligible words.

	N	Min	Max	M	SD
Arousal	392	2.40	7.86	5.2671	1.02080
Valence	392	1.40	8.33	4.6063	2.23256
Trait-state	392	1.10	2.98	1.8408	0.61003
Competence-morality	392	1.10	2.79	1.9961	0.42274

Adjectives as Means of Communicating Affect and Basis for Self-Others Evaluation

Aside from verbs and nouns, adjectives are the most important open-class words, which are present in most languages as they are used to describe persons, objects, situations, and phenomena denoted by nouns (Crystal, 2010). They also form the basis of the majority of personality questionnaires (e.g., the Big Five model itself; Goldberg, 1992; Costa and McCrae, 2008) and mood checklists, with either two-dimensional and emotional valence-based (with positive and negative affect; Watson and Tellegen, 1985), arousal-based (with energy and tension arousal; Thayer, 1989), or three-dimensional (linking the two mentioned, with energy, tension arousal, and hedonic tone; Matthews et al., 1990) being developed and used in psychology. In fact, most of the studies reporting norms for a large number of open-class words include adjectives (for an overview, see Table 1 in Riegel et al., 2015).

In Polish, as in other Indo-European languages (especially Germanic ones like English, German, and Dutch), the adjectives can be classified into five taxonomic groupings (Angleitner et al., 1990; Szarota, 1995): (1) dispositions or traits, e.g., "trustworthy," (2) temporary conditions or states, including behavioral and bodily states such as "tired," (3) social terms, i.e., roles and evaluations, e.g., "brotherly" and "unacceptable," (4) physical characteristics and appearance, e.g., "tall," and (5) terms of limited utility, i.e., technical or vague, e.g., "haphazard." The first two groupings are closely connected to the concept of affectivity. Also of paramount importance for the present study is the fact that there is a substantial number of adjectives of a "hybrid" or ambiguous kind, i.e., having both a trait and a state reading (Angleitner et al., 1990). These adjectives can be called "state-like conditions" (Ortony et al., 1987) and present a frequent yet problematic subcategory of research material. Terms such as "active," "happy," "nervous," "energetic," and "optimistic" are common examples of "state-like conditions" words. Since they seem of particular interest to the topic of affectivity, these verbal stimuli are one of the focus points of the present study.

There have been few normative word battery studies that dealt with adjectives only (Gilet et al., 2012; Ric et al., 2013; Quadflieg et al., 2014). All of them involved French words and each focused on different aspects of verbal stimuli evaluation and categorization. The study of Gilet et al. (2012) presented data on age-related ratings on the valence, arousal, and imagery dimensions of trait and state adjectives, where no distinction between state and trait was made. In the study of Ric et al. (2013), they focused on trait adjectives alone and reported evaluations of valence, approach-avoidance, and possessor-other relevance dimensions. Finally, the study of Quadflieg et al. (2014) collected a broad set of adjectives, including trait, state, and appearance ones, alongside non-human descriptors and reported ratings on the human–non-human applicability, valence, visibility, intensity, familiarity, concreteness, and temporal stability dimensions. The temporal stability seemed to be a particularly interesting dimension, ranging from very transitory (state) to very enduring (attribute). Therefore, this study found that temporal stability increased alongside applicability to non-human entities; thus, it was a marker of object (inanimate) unchangeability rather than a human affective experience.

The two aforementioned groupings of adjective classification indeed seem the most pivotal for emotional experience and communication, both interpersonal and intrapersonal. However, the social evaluation, i.e., how one judges themselves and others in terms of ability and ethics, is of great importance as far as self-worth and engagement with others are concerned (Wojciszke, 2005), and can thus be related to strictly emotional measures. These social perceptions, i.e., evaluating oneself and others, are based on competence and morality judgments and are suggested to be the most important meanings processed by laymen within the contexts of social behavior, cognition, and personality (Wojciszke, 1997). These evaluations especially relate to trait words (Ric et al., 2013). In general terms, people tend to form self-judgments based more on competence values and other judgments based on morality values (Wojciszke, 1997, 2005; Wojciszke et al., 1998). Interestingly, moral transgressions are what always elicit negative emotions in the perceiver, whereas moral acts elicit positive responses, provided the person responsible for them is liked by the observer (Wojciszke and Szymków, 2003). Thus, social-based experiences and evaluations relate to affectivity, while the specifics of this relationship within the emotional language framework are well worth investigating. How the competence and morality evaluations within the social dimension relate to emotional experience conveyed by both trait and state aspects of affectivity is another point of interest of the present study.

Affective Norms for Verbal Stimuli

To our best knowledge no previous normative word rating study combined category (trait, state), social (competence, morality) and emotional evaluations. Previous research presenting word batteries in Polish (Riegel et al., 2015; Imbir, 2016), upon which the present study is based, focused on important dimensions related to meaning and interplay between language and emotions with a broader and more general scope. The most common and well-researched factors of the affective kind, which are present in many research programs across cultures and languages, are valence (the pleasantness or unpleasantness of an object) and arousal (the internal reaction evoked by an object, ranging from calmness to extreme activation or excitement). It is worth noting that brain responses during emotional word processing have been shown to be modulated by both dimensions, either through independent or interactive manners (for a review see Citron, 2012). The origin of such an approach toward emotion-laden language could be traced to the semantic differential research by Osgood et al. (1957), in which variance in affective meaning assessments was mostly accounted for by three main dimensions: evaluation (rating something as good or bad and pleasant or unpleasant), i.e., valence, activity (active or passive and lively or still), which related closely to arousal, and potency or dominance (strong-controlling or weak-submissive).

One of the best known, and commonly used as the basis for research, normative databases is the Affective Norms for English Words (ANEW) collected by Bradley and Lang (1999), which has 1,034 words with norms for emotional valence, arousal, and dominance. Impressively, it has been extended by the study of Warriner et al. (2013) to 13,915 words. Another large affective word dataset, which includes 4,300 words, is available in Dutch (Moors et al., 2013). The Nencki Affective Word List (NAWL; Riegel et al., 2015), the Polish adaptation of the Berlin Affective Word List-Reloaded (BAWLR; Võ et al., 2009), is a list of 2,902 words rated on emotional valence, arousal, and imageability scales with control for linguistic variables (frequency, word length, and grammatical class). The study reported on 612 adjectives of various kinds, which is consistent with the groupings discussed above. The impressive Affective Norms for Polish Words Reload by Imbir (2016; ANPW_R), originally based on the ANEW list (Bradley and Lang, 1999), presented the assessments for 4,900 words on affective scales (valence, arousal, dominance, origin, and subjective significance) alongside psycholinguistic ones (concreteness, imageability, and age of acquisition) with linguistic variables control. There are 768 adjectives, again, of various groupings, which are reported there. This subset was the point of origin of the present study. In our approach, it was decided to abandon the psycholinguistic scales, i.e., the adjectives related to emotional experiences based on states and traits are rather more abstract and not readily imageable, and rather focus on the main affective ones (valence and arousal) and the essentials for the concept of affectivity categorical evaluations (state, trait, and hybrid).

Aims

The main objective of the study was to provide research materials for studies on state- and trait-based affective experiences, which would help us to better understand the bottom-up (stimulus-based) and top-down (related to goals, attitudes, and experiences; Corbetta and Shulman, 2002) processes of emotion word encoding. In a broader theoretical framework, the present work was designed to provide us with a more complete description of affectivity and, as a result, enable us to better grasp the interplay between affective dispositions (personality traits) and states (moods) of varying valence and arousal levels. State-like conditions conveyed by "hybrid" adjectives should be of particular interest in this context for researchers, especially those dealing with very sensitive and finely detailed measures of stimuli encoding, e.g., evoked potentials of the brain (EP or event-related potential ERP). The state, trait, and hybrid adjective list could prove to be beneficial especially in the line of research that tries to distinguish and detail mood-congruence and disposition-congruence effects during stimuli processing and state-trait interactions (Rusting, 1998). Because of this, and in order to build a more reliable, stable, and ready to use research base, especially for sensitive brain responses (ERP) study programs, it was decided to select the adjectives on a stricter linguistic control basis (preselection, see section Materials and Methods). Such a list could also prove useful for developing new and updating old mood checklists and personality questionnaires.

The secondary goal of the study was to explore the relations between affective (and category) adjective norms and social judgment norms based on competence and morality evaluations. These evaluations form the basis for self- and other-perceptions and have probable, yet not clearly established, connections with the affective experience.

Similar to the case with other normative studies in Polish (Riegel et al., 2015; Imbir, 2016), it was predicted in this study that the dataset obtained would be reliable (split-half estimates) and stable (correlations with other databases, especially with the ratings of the ANPW_R, since the present study shared a bulk of adjectives with the set). The study also predicted that there will be a substantial number of adjectives categorized as hybrids, In an exploratory fashion, we also expected the social judgment ratings to correlate with other scales (possibly most notably with the affective ones, especially valence).

MATERIALS AND METHODS

Participants

One hundred and ninety-five young adults who were students of Jagiellonian University with an *M* age of 22.36, $SD = 2.23$ participated in the study. Most of the participants were women ($N = 150$, *M* age 22.21, $SD = 2.08$), with a minority of men ($N = 45$, *M* age 22.87, $SD = 2.63$). The vast majority of participants were right-handed with only 10 participants declaring left-handedness (women $N = 9$, men $N = 1$). All of the participants signed a consent form before the procedure. After they completed the ratings, they were paid a Polish currency equivalent of €3.5 for their participation.

Study Design and Materials

A special effort was made to prepare a more robust adjective list from those present in Polish word datasets (Riegel et al., 2015; Imbir, 2016), so that we could select the most representative and linguistically matched cases (preselection) from them. Firstly, all the adjectives from the larger ANPW_R (Imbir, 2016) list were extracted, then the list was supplemented with the adjectives from the NAWL (Riegel et al., 2015), which were not present on the ANPW_R. Afterward, two judges, one holding a Ph.D. in psychology and the second a Ph.D. in Polish literature, selected all of the adjectives denoting and connoting traits and states relating to human affective experience, in essence, personality, and mood adjectives from the compiled list. This yielded a list of 401 adjectives. Then, we extracted all the adjectives from the mood checklists (Watson and Tellegen, 1985; Thayer, 1989; Matthews et al., 1990) and the personality questionnaires and lexical studies on personality (Cattell, 1943; Goldberg, 1992; Costa and McCrae, 2008; see also De Raad, 2000) with translations based on the Polish adaptations of the tools and the individual differences manual of Strelau (Szarota, 1995; Strelau, 2014). Finally, the judges went through a small contemporaneous dictionary of the Polish language (Sobol, 2006) and selected from it all the mood and personality adjectives. The complete list comprised 1,061 adjectives. Linguistic variables for stricter control were obtained for all the adjectives: frequency values from two datasets, as was the case in the study of Imbir (2016), Subtlex_pl film and TV subtitle database (Mandera et al., 2015) and literature and electronic texts collection by Kazojć (2011) alongside word length values, i.e., letter count. Based on those values, 400 adjectives were then selected for normative ratings. The final list comprised the most average words, located effectively within 0.75 SD of the respective means[1].

Procedure

The participants were divided into four groups, with each group required to rate 100 adjectives. The first group consisted of 48 participants (women $N = 32$), the second 51 (women $N = 38$), the third 48 (women $N = 42$), and the fourth 48 (women $N = 39$). All the groups rated the adjectives on the same four dimensions. Two of these dimensions were affective and exactly the same as seen in the study of Imbir (2016). Both were also based on the Self-Assessment Manikin (SAM) used in the ANEW ratings (Bradley and Lang, 1999) and used nine-point scales. The emotional valence dimension rating was between 1 (most negative) and 9 (most positive) and the emotional arousal was from 1 (completely non-arousing) to 9 (extremely arousing). The third dimension was a category one rated on a three-point scale, where 1 meant that the shown adjective denoted a stable trait, 3 a transient state, and 2 that it could denote both a trait and a state. The last dimension was a social judgment one, and it was constructed, respectively, in the same fashion as the category one, with 1 meaning that the shown adjectives denoted a competence evaluation, 3 a moral one, and 2 both competence and moral judgment. **Figure 1** shows a sample word with the rating scales as seen by the participants during the rating session, alongside the description of the scales presented to the participants at the beginning.

After the participants arrived in the laboratory, they were seated comfortably in a separate room. Then, it was explained to them that they were to rate 100 words and were further asked to try to pay close attention to each word and not dwell too much on the answer. They were also warned that each rating was final and that it would not be possible to go back and answer again. The participants then signed the consent forms and the rating procedure was run in the PsychoPy2 software (University of Nottingham) (Peirce et al., 2019), where detailed instructions with the scales description were presented at the start. The participants read the instructions at their own pace and were encouraged to ask questions if anything was

[1]The frequency values of the adjectives qualified for the ratings were between 1 and 669 in Subtlex_pl dataset and 1 and 768 in the collection of Kazojć. For a word to be excluded from the final set, it had to have a frequency below 1 and above 1,100 on just one dataset or exceed the upper frequency value on both datasets. The word length of the adjectives qualified was between 7 and 11 letters.

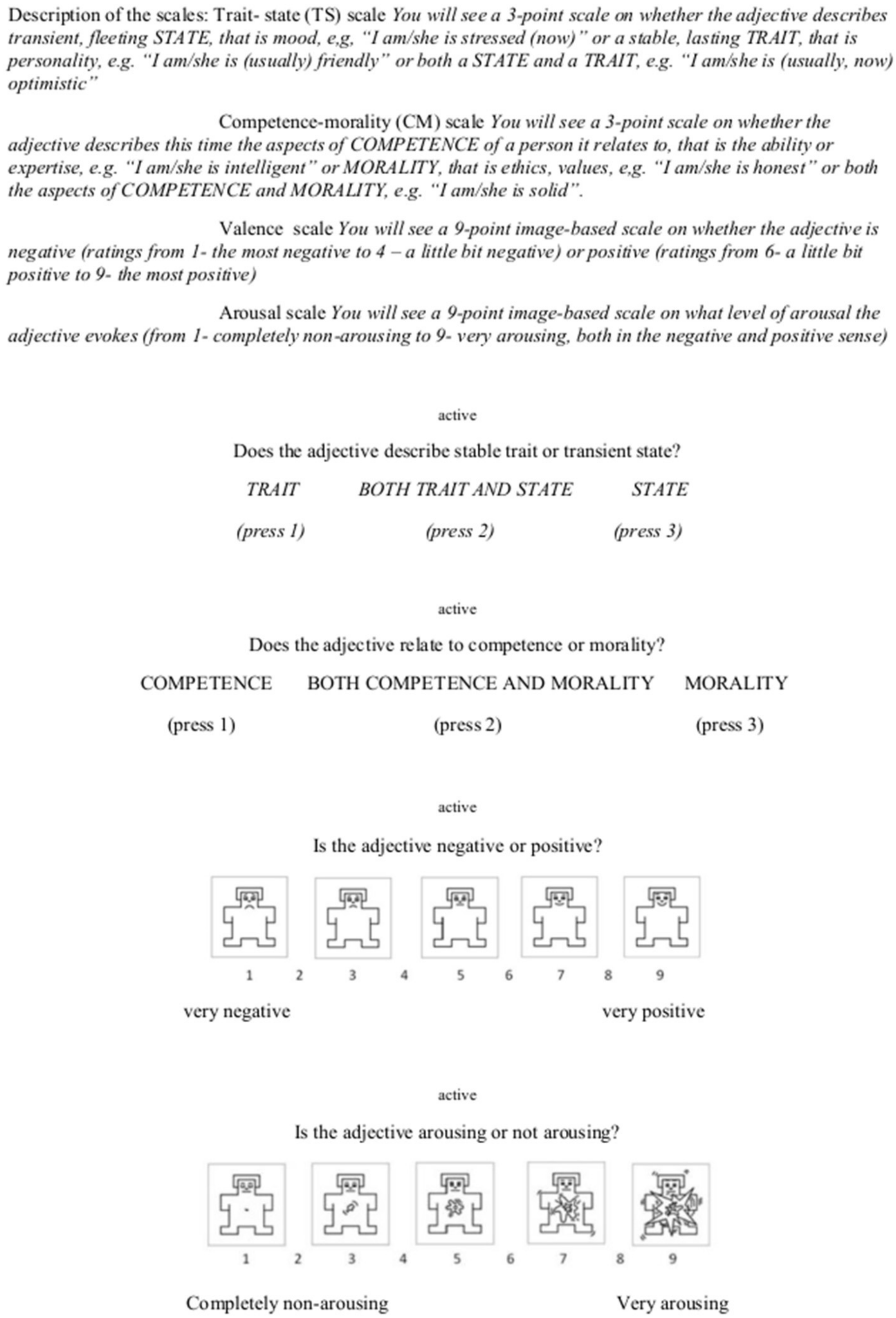

FIGURE 1 | Descriptions of the scales seen during the introductory information period of the rating session and subsequent sample rating screens with each scale as presented to the participants. Self-Assessment Manikin (SAM) based on (Bradley and Lang, 1999).

unclear. After the instructions were read, two sample words with full scales were presented; the samples were not a part of the 400 adjectives to be rated. Then, the actual rating session began. The participants rated each word on four scales by pressing the numerical keys on a keyboard. Each adjective was presented on the screen in a central position for 2 s, followed by the rating scales appearing, each one at a time, with the adjective to be rated visible in smaller fonts above the scale (see **Figure 1**). The next scale appeared immediately after the rating key was pressed. On average, it took 30 min to complete the 100-word rating. In the end, the participants were presented with a sheet of paper on which all 100 adjectives were printed and asked to mark the words that they did not comprehend. After that, the participants were thanked, paid, and were free to leave.

The procedure was prepared in two versions, namely, for women and men because Polish is a fusional language and has suffixes pointing to the gender of the nouns. Thus, an adjective "active," "aktywny" in Polish, is a male form of the word (because of the suffix "y"), while the female form is "aktywna." Afterward, all the 400 adjectives were presented in one form depending on the sex of the participant.

RESULTS

All statistical analyses were conducted in the IBM SPSS 26 software, New York, USA. In the beginning, descriptive statistics, i.e., *M* and *SD*s, for every adjective and rating scale [valence, arousal, trait-state (TS), and competence-morality (CM)] were obtained. The complete list of 400 adjectives (original and translated into English) with the statistics (alongside linguistic values for frequency and letter count) is available as **Supplementary Material** to the study. Afterward, the number of adjectives described as not understood at the end of the rating session was analyzed, and all the words marked as unintelligible by more than 25% of the participants who rated them were excluded from further analyses. The words that were not well-understood were: "afektowany" ("mincing"), "egzaltowany" ("exalted"), "lubieżny" ("lecherous"), "obcesowy" ("objectionable"), "sterany" (one of the synonymous variants of "weary"), "występny" ("debased"), "zapiekły" ("fiery"), and "zdrożony" (another of the synonymous variants of "weary")[2]. **Table 1** presents the general descriptions for the included 392 adjectives. In order to test for the normal distribution of the ratings, Shapiro–Wilk tests were used. All the rated dimensions were shown to have non-normal distributions. **Figure 2** shows the distributions of the four dimensions ratings. Emotional valence and trait-state were bimodal (with TS being clearly negatively biased toward the lower end of the scale), whereas the distribution of CM and arousal only resembled normal distribution (with CM being centered close to the midline and arousal being slightly positively biased toward the high end of the scale).

Each dimension was divided into low, moderate, and high or one category or mixed type based on the rating ranges of average scores. The emotional valence dimension was divided into low-valence (negatively rated) adjectives (ratings ranging from 1 to 3.99), moderate (neutral) adjectives (ratings from 4 to 6), and high-valence (positive) adjectives (ratings from 6.01 to 9). The criteria for this classification were based on previous studies on normed word datasets (Ferré et al., 2012; Monnier and Syssau, 2014; Imbir, 2016), which used the same SAM rating scales and divided the words similarly. Similarly, the likewise scaled emotional arousal dimension was divided into adjectives that were low (ratings from 1–3.99), moderate (4–6), and high (6.01–9) in arousal. The dimensions TS and CM were evaluated on simpler three-point scales, in which adjectives with the lowest ratings (from 1 to 1.49) were classified as trait and competence ones, adjectives with moderate ratings (from 1.5 to 2.5) were hybrid (TS) or mixed (CM), and the adjectives with the highest scores (1.51–3) were state and morality words, respectively. The justification for such ranges was that the single category evaluations, e.g., trait or state, are simple and obvious, meaning the categorical denotation of the word is easily grasped and so the population scores should be located close to the polar ends of the scales, i.e., low or high. In contrast, the hybrid and mixed category evaluations have a level of uncertainty *vis-à-vis* having a broader denotative and added connotative meaning, so their classification should be based on broader average scores across the tested population. The number of adjectives classified on the subdivided dimensions based on average rating score ranges is presented in **Table 2**.

Reliability and stability analyses were conducted to assess the validity of the ratings. Finally, to verify the relations between the dimension correlation analyses were conducted, with an additional and tentative analysis of the differences between sexes at the end.

Reliability

The reliability of the assessment was measured by applying the split-half estimation across the whole sample. The split was based on the odd and even numbering of the participants, with special care taken with respect to the less frequent men in order to balance their number in each half. The mean rating for each adjective in each group was then calculated and Pearson correlations were applied within each dimension. Because of the smaller number of participants in the split-half comparisons as compared to the whole set, the correlations were then adjusted through the use of the Spearman–Brown formula. All the correlations were strong ($r > 0.9$) and significant ($p < 0.01$). **Table 3** presents the results.

Stability

To measure the stability of the ratings, *r*-Pearson correlations were applied for the 132 adjectives that were qualified for the final 400-word set and that were also a part of the ANPW_R dataset of Imbir (2016). Since only two scales (emotional valence and arousal) were exactly the same as what was used in the work of Imbir, those two affective dimensions were correlated. Both correlations were significant ($p < 0.01$), but in particular, the emotional valence one was especially high ($r = 0.916$), whilst arousal was moderately strong ($r = 0.615$). To further validate the ratings, we additionally correlated word scales from our dataset with the adjectives from the English-based ANEW dataset (Bradley and Lang, 1999) and the biggest French-based adjective normed study (Quadflieg et al., 2014). This was achieved by matching the English translations of the adjectives. Correlations with the ANEW datasets, based on the 62 adjectives found in both datasets, were significant ($p < 0.01$) and high; again, especially for the emotional valence dimension ($r = 0.855$) and the arousal dimension, but to a lesser extent ($r = 0.68$). The same was true for the correlations with the French study (all significance levels

[2]Interestingly, but not surprisingly, the absolutely least understood word was "mincing," which was described as incomprehensible by 81% of the participants rating the word.

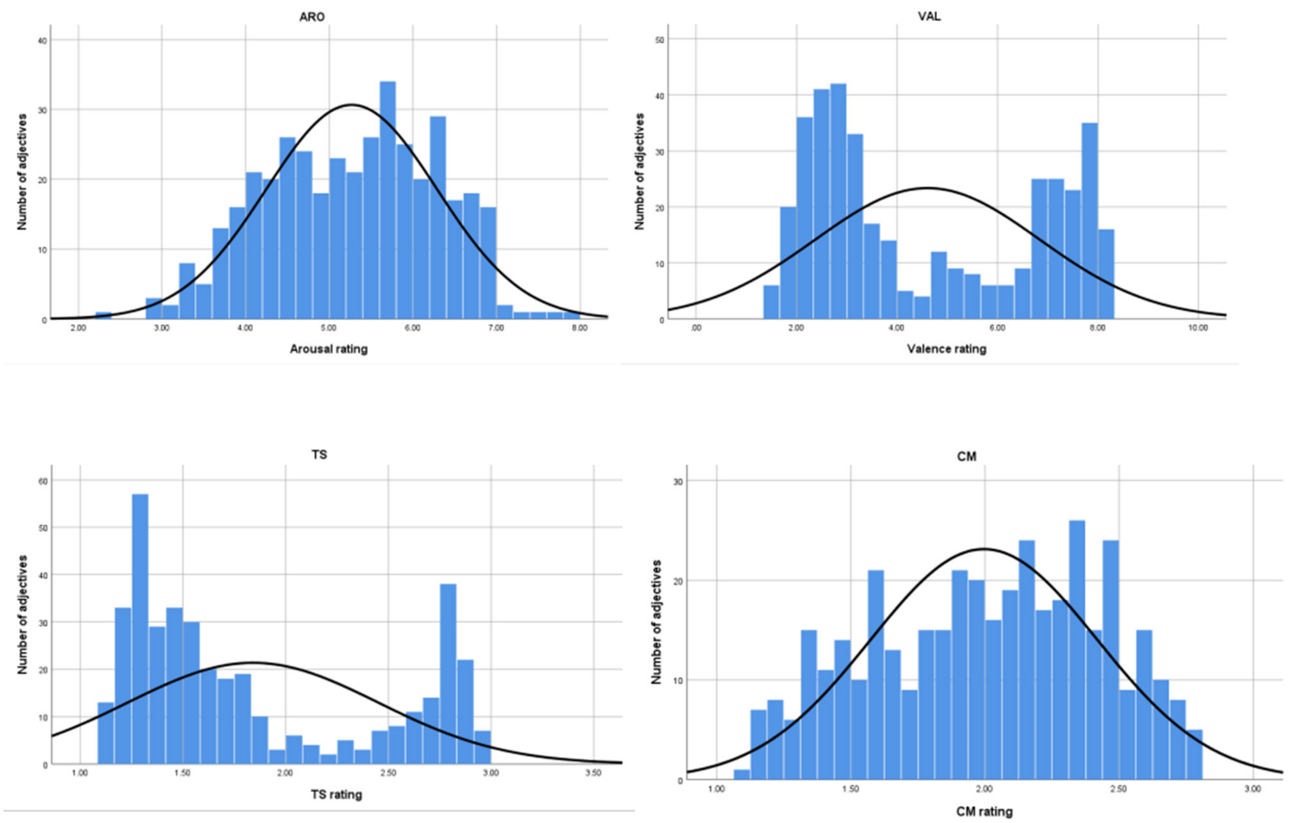

FIGURE 2 | Histograms presenting the number of adjectives rated in one-point intervals of the scales of the respective dimensions' (ARO, arousal; VAL, valence; TS, trait-state; CM, competence-morality).

$p < 0.01$), where, again, 132 matching adjectives were found, and we compared three dimensions: emotional valence ($r = 0.919$), emotional arousal, called "intensity" in the referenced dataset ($r = 0.388$), and temporal stability, related to the trait-state dimension of our study, only with reversed scaling ($r = -0.760$). All of the results related to the stability measures can also be seen in **Table 3**.

Correlations Between Variables

In order to check for relations between the measures, *r*-Pearson correlations were applied for all the dimensions, i.e., affective, category, social judgment alongside linguistic variables). Because of the non-normal distribution of the variables, we did not conduct a partial correlation analysis. **Table 4** presents all the results. Here only the most pertinent, significant, and strongest correlations are described.

The emotional valence dimension was found to be negatively correlated weakly and moderately with the TS ($r = -0.23$) and CM dimensions ($r = -0.305$), respectively, whilst arousal was weakly positively correlated with competence-morality ($r = 0.25$). **Figure 3** shows the scatterplots of the correlated ratings. The emotional valence correlation with the category dimension would tentatively point to the fact that positive adjectives are categorized rather as trait ones, whereas valence relation to social judgement would mean that negative adjectives are judged as more related to morality. The correlation between emotional arousal and CM dimensions would, in turn, point to the observation that the more arousing adjectives are morality ones.

Additionally, comparisons between sexes were conducted (although one has to note the uneven number of male and female

TABLE 2 | Numbers and percentages of the adjectives rated on every dimension (divided into low, moderate, and high or one category type or a mixed category based on the rating ranges).

	Number of adjectives	% of the included adjectives (392)
Arousal low	48	12
Arousal moderate	242	62
Arousal high	102	26
Valence low (negative)	209	53.3
Valence moderate (neutral)	45	11.5
Valence high (positive)	138	35.2
Trait	165	42
Trait-state hybrid	127	32.5
State	100	25.5
Competence	62	16
Competence-morality mixed	284	72
Morality	46	12

TABLE 3 | Reliability and stability estimations for every dimension.

	Split-half correlation (r-Pearson's)	Split-half correlation Spearman-Brown adjustment	Correlations with the ANPW_R	Correlations with the ANEW	Correlations with Quadflieg et al. (2014)
Arousal	0.905	0.950	0.615	0.680	0.388
Valence	0.986	0.993	0.916	0.855	0.919
Trait-state	0.969	0.984	–	–	–0.760
Competence-morality	0.910	0.953	–	–	–

All $p < 0.01$.

TABLE 4 | r-Pearson correlations between the variables.

	Subtlex_pl frequency	Kazojć (2011) frequency	Number of letters	Valence	Arousal	TS	CM
Subtlex_pl frequency	–	0.512**	–0.077	0.025	0.176**	–0.079	0.136**
Kazojć (2011) frequency	0.512**	–	–0.178**	0.118*	0.022	0.141**	0.056
Number of Letters	–0.077	–0.178**	–	–0.076	0.087	0.070	0.037
Valence	0.025	0.118*	–0.076	–	0.075	–0.230**	–0.305**
Arousal	0.176**	0.022	0.087	0.075	–	–0.076	0.250**
TS (trait-state)	–0.079	0.141**	0.070	–0.230**	–0.076	–	0.014
CM (competence-morality)	0.136**	0.056	0.037	–0.305**	0.250**	0.014	–

**$p = 0.05$; ***$p = 0.01$.*

participants, so all the differences and relations reported here should be seen as tentative at best). The r-Pearson correlations of all the dimensions for all the included words as rated by female and male participants were applied. All of the rated dimensions were significantly and strongly positively correlated between the sexes: arousal $r = 0.682$ ($p < 0.01$), valence $r = 0.958$ ($p < 0.01$), TS $r = 0.905$ ($p < 0.01$), CM $r = 0.789$ ($p < 0.01$). The possible differences between the ratings were verified by paired *t*-tests. However, only the arousal dimension ratings proved to be significant, $t_{(391)} = 3.77$ ($p < 0.001$), with the adjectives being rated as more arousing by women (mean rating 5.29, SD = 1.29) than men (mean rating 5.11, SD = 1.14).

DISCUSSION

In a broader theoretical framework, the study was designed to describe in detail the linguistic means of encoding and communicating emotional experiences related closely to the mood-dispositional concept of affectivity (Watson and Clark, 1984; Tellegen, 1985). It can also be viewed as supplementary material to the two broad Polish word normative datasets, the ANPW_R (Imbir, 2016) and the NAWL (Riegel et al., 2015), focusing on one grammatical class of characteristics, which is particularly interesting for affective psychology and neuroscience, i.e., the adjectives denoting and connoting human traits and states. The dataset presented in the current study should be viewed as more closely related to the ANPW_R list since it shares 33% of the adjectives with it, two of the main affective scales (valence, arousal), and the linguistic measures (frequency, word length) used as the basis for the ratings.

Affectivity may be a very useful concept since it better orients and nuances our knowledge on the relationship between transient, state-like and stable, and trait-based emotional experiences. In terms of PA and NA, they also have applied values, whereas it is the former that can be successfully used as a diagnosis and prediction tool for depression (Watson et al., 1988) while both PA and NA can predict job performance (Kaplan et al., 2009). The analyses based on these observations could also prove useful for broader inferences on language and culture (Wierzbicka, 1994). In the case of the present study, the proportion analysis of the affective and category evaluations have shown that more than half of the selected 400 adjectives were of negative valence (53%) and nearly half (42%) denoted and connoted stable dispositions (traits). It is worth mentioning again that the 400 adjectives were the most average subset of a broad and representative sample of all the state and trait adjectives used in language and psychological research (see section Study Design and Materials). Future psycholinguistic and cross-cultural research could determine if such a pattern of relevance and importance embedded in the language (lexical hypothesis) focused more on the negative and fixed affective phenomena is indeed a real cultural characteristic, along with other patterns under what conditions that could possibly be observed.

In terms of a more tangible and immediate benefit, the dataset of the current study could prove useful for complex research into the mood or trait congruence aspects of emotional stimuli processing (Rusting, 1998). Since the adjectives were preselected on the most-representative basis of linguistic variables (double-sourced frequency values and word length values), the set could be especially well-suited for, e.g., pilot and main studies focused on sensitive stimuli processing measures, thus necessitating very

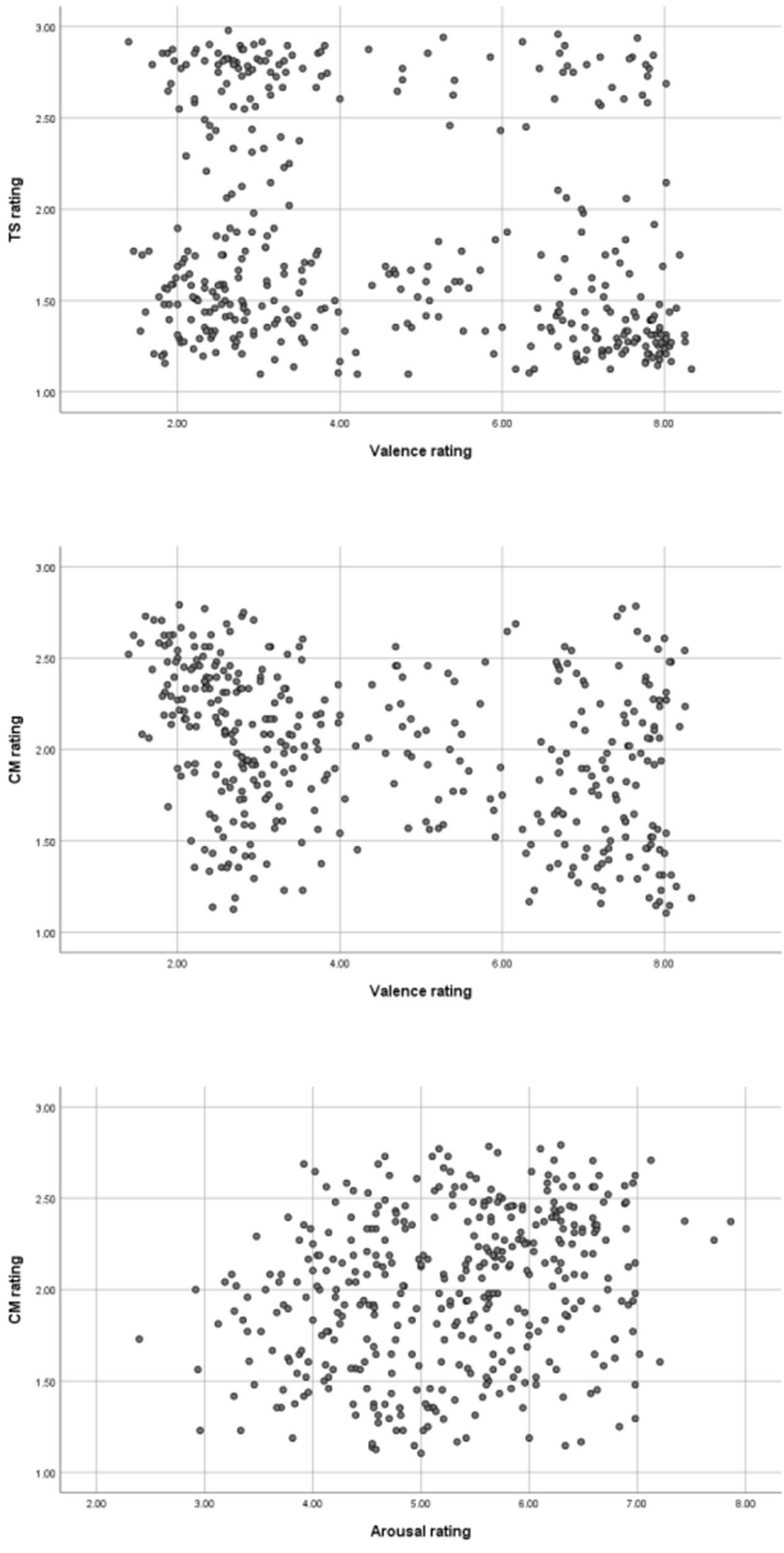

FIGURE 3 | Scatterplots of the significant correlations between affective (valence and arousal), category (TS, trait-state), and social judgment (CM, competence-morality) dimension ratings.

strict matching and control of the independent variable, i.e., brain activity research programs, especially those using the evoked-potential technique.

In order to test the stability and reliability of the dataset, we used analyses similar to those reported in other Polish word norming studies (Riegel et al., 2015; Imbir, 2016). Although the present adjective-only study was much narrower in scope, it still required these kinds of analyses to properly validate it and establish its appropriateness and usefulness for further research. In fact, other adjective-based normed studies also measured the datasets for the reliability or both reliability and stability (Gilet et al., 2012; Quadflieg et al., 2014). We predicted that the measures would prove stable and reliable. This was indeed the case. The split-half assessment showed the dataset to be very reliable, as all the correlations noted in the analysis were very strong indeed. Based on the correlations with the same dimensions as seen in the ANPW_R, on the basis of a significant portion of the adjectives rated in the current study, the stability assessment proved to be high and satisfactory. However, it is worth noting that this was based only on two affective scales and not all of the 400 adjectives that were part of the estimation (in fact, it was 33% of the whole set). Also, the correlation score of the emotional arousal dimension was only moderately strong, which was probably due to the distribution of the arousal rating in the work of Imbir (2016), which had slightly different characteristics. Although also approximately normal in nature, it was clearly negatively biased (toward the less arousing ratings), whereas the distribution of arousal in the current study was slightly positively biased. Overall, the ANPW_R set consisted of less arousing stimuli. The same is true for the additional stability measures based on correlated data with English- and French-based datasets, which further validated our ratings (see quite similar correlations obtained with the ANEW to those with the ANPW_R, **Table 3**).

The French study (Quadflieg et al., 2014) comparison is especially interesting in this context. As in the case of the current research program, its database consisted of adjectives only and the number of matching adjectives between the two datasets was substantial ($n = 132$). The emotional valence correlation was high, as expected, whereas the arousal correlation, albeit still significant, was lower, mainly because of the lower average score (less arousing word set as a whole) and the dimension of intensity used in the French study carrying not quite the same denotations and connotations (how "extreme" and not "arousing" the meaning of each word was). Interestingly, we also correlated another similar dimension present in the dataset of the study of Quadflieg et al. (2014), namely, temporal stability, (mentioned in the introduction section) with the TS dimension of the current study, with the correlation proving to be high and significant. It was also negative, since stable and enduring dispositions (traits and attributes of the two scales, respectively) were scored on the low end of the TS and the high end of the temporal stability scale. This proves that the TS scale is an important and sensitive dimension of adjective-based semantics and that the transient vs. enduring aspect of concepts expressed through language is prominent and more general in nature since it is present in words relating to human experiences and dispositions and in those relating to the attributes of inanimate objects.

As a side note, most studies cross-validating their ratings with other datasets have found that emotional valence correlations are stronger than arousal correlations (Warriner et al., 2013; Riegel et al., 2015; Imbir, 2016). This was clearly the case in the present study (see **Table 3**). Emotional valence seems to generalize very well across languages and research programs, whilst arousal shows more variability, possibly due to its stronger dependence on situational aspects, the broader connotations it evokes (which may provoke "the lost in translation" effects across languages), and its relation to affective states (the intensity of the actual feelings during the rating sessions).

This study also predicted that the category evaluation would reveal a substantial number of hybrid adjectives, i.e., denoting and connoting both trait- and state-based experiences, depending on the reading, which could be of high practical value in the language; further providing possible confounds in research (see above Rusting, 1998). Indeed, 32.5% of all the (comprehensible) adjectives rated were of this kind. This means that the hybrid adjectives were the second most abundant type of words related to affectivity (trait ones being the most common). Since our research material was sourced in various ways, one of which was a dictionary study, the proportion observed here seems like a good representation of the contemporary Polish language content. Because of this, and making use of the two original evaluations (category and social judgment), the current study is differentiated from other adjective-only based normed datasets (Gilet et al., 2012; Ric et al., 2013; Quadflieg et al., 2014) and, based on the category evaluation, it can be of applied value as a source of essential material for new or updated personality questionnaires and mood checklists.

As is the case with all the word normative research, we analyzed the relations between the measures used in the study. First of all, we did not find a significant correlation between the affective dimensions (valence and arousal), which is surprising, since it is usually strongly evidenced in most of the previous word rating research (Gilet et al., 2012; Quadflieg et al., 2014; Riegel et al., 2015; Imbir, 2016). This correlation typically points to a U-shaped relationship between emotional valence and arousal (Bradley and Lang, 1999; Moors et al., 2013; Warriner et al., 2013), meaning the more extreme the valence rating (either positive or negative), the higher the arousal score of a given stimulus. The words neutral in their emotional valence are typically rated moderately or low on their arousal. However, the lack of correlation in the present study can be explained by the nature of the preselected word set. We based the aims of the study and our search for the materials on the affectivity concept, which, at its core, has two orthogonal dimensions of positive and negative affect (Watson and Tellegen, 1985). Those dimensions are varied in arousal within their poles. Therefore, both high scores on the valence scale (positive evaluations) and low scores (negative evaluations) dominating in our dataset (there are few neutral words relating to personality and mood in the language) relate to words of moderate to strong arousal ratings (see the numbers of arousal-rated adjectives in **Table 2** and the distributions in **Figure 2**). In other words, the emotional valence dimension is clearly bimodal, with very few neutral cases, whilst the arousal dimension approximates normal distribution, with a mild positive bias. In such a clear-cut manner, this was not the case in other normative studies for vast arrays of words of different grammatical classes, including the ANPW_R (Imbir, 2016).

This study found weak to moderate correlations between the category and emotional valence dimensions and between the social judgment and valence dimensions, the latter of which was predicted based on research in social and personality psychology (Wojciszke et al., 1998; Wojciszke and Szymków, 2003). It was also the strongest observed (pertinent) correlation between measures (albeit still moderate at best, see **Table 4**)[3]. In essence, it points to a relationship in which the negative adjectives are rated as rather morality-based (see scatterplots, **Figure 3**). This is very interesting in the context of self–other perception and self–other evaluation (impressions and emotions), in which one tends to judge oneself on the competence basis and others on the morality basis (Wojciszke, 2005). If the negative adjectives embedded in language tend to be more related to morality aspects of social judgments, then it could follow that one has a (bottom-up and lexicon-based) bias toward evaluating others in a negative light. If the converse is true, then we speak mostly positively about ourselves, which could be one of the reasons behind a known characteristic of healthy people, namely, the fact that they declare feeling good and happy about themselves and their lives (Diener and Diener, 1996). These fascinating interconnected aspects of language, emotion, and social evaluations certainly merit further research. However, one of the limitations of the study must be mentioned. The three-point scale of the competence-morality dimension may be the source of some confusion and probably "flattened" the richness of this aspect. The scale was conceived as a counterpart to the category three-point scale, which seems to be a good enough measure for the purposes of the current study (see the numbers of the category ratings in **Table 2**). We wanted to avoid adding another, differently balanced scale, e.g., a four-point scale with an additional position, "neither" competence nor morality, to not tire and unnecessarily confuse the participants with different numbering (key press) systems. This was an oversimplification that might have worked disadvantageously for the study, thus, affecting the results, since the vast majority of the adjectives rated on the scale were evaluated as both competence and morality ones. Some of the participants reported at the end of the rating session that they picked this middle response because there was no response labeled "neither." Such an option would be beneficial for any rating scale based on intricate evaluations, and social judgement is no exception. On the other hand, "neither" or "do not know" answers present in the scale could result in an "easy way out" for the tired participants rating dozens of words, which could further distort the evaluations. Therefore, 5-, 7-, or even 9-point scales seem like better solutions for a nuanced and complex dimension such as social judgments. This limitation must be strongly considered when interpreting all the aspects related to the CM measure.

[3]The strongest correlation was the obvious one, of interest only strictly linguistically, between two frequency measure subsets.

Additionally, we looked for tentative relations in ratings between participants who were women and men, since the word norms for both sexes, although correlated, tended to differ especially *vis-à-vis* the affective variables (Riegel et al., 2015; Imbir, 2016). This study applied only the simplest of solutions and, thus, urge caution in interpreting the data since the groups of women and men were uneven in number. Sex-based ratings on all the dimensions were similar and highly positively correlated. Only the emotional arousal correlation was of moderate strength. By comparing the means of the ratings, we found that, indeed, the arousal rating differed significantly between the sexes, with higher ratings on average for all the adjectives evaluated by participants who were women. It could be a slight encouragement for further verification of the possibility that women are more sensitive to the emotional (arousing) content of words and, thus, evaluated them differently based on that variable. However, the picture is certainly more complex and possibly based on the interactive effects of at least valence and arousal in between sex comparisons (Riegel et al., 2015). The uneven woman–man number of participants is another limitation of the study. Although the ratings were similar and positively correlated to a high degree (and so the male ratings were included in the general dataset), the list should be viewed as more representative for the women population.

Lastly, the overall number of the adjectives included in the database ($n = 400$) may seem small, especially when one considers using words of this grammatical class in complex research programs, which require matching the words on several psycholinguistic conditions or filling the given sub-category with several dozens of stimuli for comparisons. Four hundred may be too small a number to begin with when considering the satisfaction of all the conditions. On the other hand, one of the presumed strengths of this dataset is its compactness and "pre-matching" of the linguistic variables that influence greatly early stages of stimuli encoding, namely, word length and frequency, which should make devising research programs easier.

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/s.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

SG contributed to the study design, the analysis and interpretation of data, and the drafting and critical revision of the manuscript. MW contributed to the study design and the gathering of the data. HC contributed to the gathering of the data and the critical revision of the manuscript. AT contributed to the gathering of the data. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by a grant from the National Science Centre, Poland, 2016/23/D/HS6/01645.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2021.683012/full#supplementary-material

Supplementary Table 1 | List of 400 adjectives and their ratings on valence, arousal, category, and social judgments dimensions alongside linguistic (frequency and word length) measures.

REFERENCES

Angleitner, A., Ostendorf, F., and John, O. P. (1990). Towards a taxonomy of personality descriptors in German: a psycho-lexical study. *Eur. J. Pers.* 4, 89–118.

Bradley, M. M., and Lang, P. J. (1999). *Affective Norms for English words (ANEW): Stimuli, Instruction Manual, and Affective Ratings*. Technical report C-1, The Center for Research in Psychophysiology, University of Florida, Gainesville, FL.

Cattell, R. B. (1943). The description of personality: basic traits resolved into clusters. *J. Abnormal Soc. Psychol.* 38, 476–506. doi: 10.1037/h0054116

Citron, F. M. M. (2012). Neural correlates of written emotion word processing: a review of recent electrophysiological and hemodynamic neuroimaging studies. *Brain Lang.* 122, 211–226. doi: 10.1016/j.bandl.2011.12.007

Corbetta, M., and Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3, 201–215. doi: 10.1038/nrn755

Costa, P. T. Jr., and McCrae, R. R. (2008). "The revised NEO personality inventory (NEO-PI-R)," in *The SAGE Handbook of Personality Theory and Assessment, Vol 2: Personality Measurement and Testing,* eds G. J. Boyle, G. Matthews and D. H. Saklofske (London: Sage Publications, Inc.), 179–198. doi: 10.4135/9781849200479.n9

Crystal, D. (2010). *The Cambridge Encyclopedia of Language*. Cambridge: Cambridge University Press.

De Raad, B. (2000). *The Big Five Personality Factors: The Psycholexical Approach to Personality*. Gottingen: Hogrefe and Huber Publishers.

Diener, E., and Diener, C. (1996). Most people are happy. *Psychol. Sci.* 7, 181–185. doi: 10.1111/j.1467-9280.1996.tb00354.x

Ferré, P., Guasch, M., Moldovan, C., and Sánchez-Casas, R. (2012). Affective norms for 380 Spanish words belonging to three different semantic categories. *Behav. Res. Methods* 44, 395–403. doi: 10.3758/s13428-011-0165-x

Gilet, A.-L., Grühn, D., Studer, J., and Labouvie-Vief, G. (2012). Valence, arousal, and imagery ratings for 835 French attributes by young, middle-aged, and older adults: the French Emotional Evaluation List (FEEL). *Eur. Rev. Appl. Psychol.* 62, 173–181. doi: 10.1016/j.erap.2012.03.003

Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychol. Assess.* 4, 26–42. doi: 10.1037/1040-3590.4.1.26

Imbir, K. K. (2016). Affective norms for 4900 Polish words reload (ANPW_R): assessments for valence, arousal, dominance, origin, significance, concreteness, imageability and, age of acquisition. *Front. Psychol.* 7:1081. doi: 10.3389/fpsyg.2016.01081

Kaplan, S., Bradley, J. C., Luchman, J. N., and Haynes, D. (2009). On the role of positive and negative affectivity in job performance: a meta-analytic investigation. *J. Appl. Psychol.* 94, 162–176. doi: 10.1037/a0013115

Kazojć, J. (2011). *Słownik Frekwencyjny Jezyka Polskiego [Polish Language Frequency Dictionary]*. Available online at: https://doczz.pl/doc/911406/s\T1\lownik-frekwencyjny-jezyka-polskiego-v.05.2011

Mandera, P., Keuleers, E., Wodniecka, Z., and Brysbaert, M. (2015). Subtlex-pl: subtitle-based word frequency estimates for Polish. *Behav. Res. Methods* 47, 471–483. doi: 10.3758/s13428-014-0489-4

Matthews, G., Jones, D. M., and Chamberlain, A. G. (1990). Refining the measurement of mood: the UWIST Mood Adjective Checklist. *Brit. J. Psychol.* 81, 17–42. doi: 10.1111/j.2044-8295.1990.tb02343.x

Monnier, C., and Syssau, A. (2014). Affective norms for french words (FAN). *Behav. Res. Methods* 46, 1128–1137. doi: 10.3758/s13428-013-0431-1

Moors, A., De Houwer, J., Hermans, D., Wanmaker, S., van Schie, K., Van Harmelen, A.-L., et al. (2013). Norms of valence, arousal, dominance, and age of acquisition for 4,300 Dutch words. *Behav. Res. Methods* 45, 169–177. doi: 10.3758/s13428-012-0243-8

Ortony, A., Clore, G. L., and Foss, M. A. (1987). The referential structure of the affective lexicon. *Cogn. Sci.* 11, 341–364. doi: 10.1016/S0364-0213(87)80010-1

Osgood, C. E., Suci, G. J., and Tannenbaum, P. (1957). *The Measurement of Meaning*. (Chicago, IL: University of Illinois Press).

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., et al. (2019). PsychoPy2: experiments in behavior made easy. *Behav. Res. Methods* 51, 195–203. doi: 10.3758/s13428-018-01193-y

Quadflieg, S., Michel, C., Bukowski, H., and Samson, D. (2014). A database of psycholinguistic and lexical properties for French adjectives referring to human and/or nonhuman attributes. *Can. J. Exp. Psychol.* 68, 67–76. doi: 10.1037/cep0000001

Ric, F., Alexopoulos, T., Muller, D., and Aubé, B. (2013). Emotional norms for 524 French personality trait words. *Behav. Res. Methods* 45, 414–421. doi: 10.3758/s13428-012-0276-z

Riegel, M., Wierzba, M., Wypych, M., Zurawski, Ł., Jednoróg, K., Grabowska, A., et al. (2015). Nencki affective word list (NAWL): the cultural adaptation of the Berlin Affective Word List–Reloaded (BAWL-R) for Polish. *Behav. Res. Methods* 47, 1222–1236. doi: 10.3758/s13428-014-0552-1

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychol. Rev.* 110, 145–172. doi: 10.1037/0033-295x.110.1.145

Russell, J. A. (2009). Emotion, core affect, and psychological construction. *Cogn. Emot.* 23, 1259–1283. doi: 10.1080/02699930902809375

Rusting, C. L. (1998). Personality, mood, and cognitive processing of emotional information: Three conceptual frameworks. *Psychol. Bull.* 124, 165–196. doi: 10.1037/0033-2909.124.2.165

Sobol, E. (2006). *Słownik Jezyka Polskiego PWN [Polish Dictionary]*. Wydawnictwo Naukowe PWN.

Strelau, J. (2014). *Róznice Indywidualne. Historia – Determinanty – Zastosowania*. Scholar. Available online at: https://scholar.com.pl/en/books/1528-roznice-indywidualne-historia---determinanty---zastosowania-wydanie-drugie.html

Szarota, P. (1995). Polska Lista Przymiotnikowa (PLP): Narzedzie do diagnozy Pieciu Wielkich czynników osobowości. [Polish Adjective List: an instrument to assess the five-factor model of personality.]. *Studia Psychol.* 33, 227–256.

Tellegen, A. (1985). "Structures of mood and personality and their relevance to assessing anxiety, with an emphasis on self-report," in *Anxiety and the Anxiety Disorders,* eds A. H. Tuma and J. Maser (New York, NY: Lawrence Erlbaum Associates, Inc.), 681–706.

Tellegen, A., and Waller, N. G. (2008). "Exploring personality through test construction: development of the multidimensional personality questionnaire," in *The SAGE Handbook of Personality Theory and Assessment, Vol 2: Personality Measurement and Testing,* eds G. J. Boyle, G. Matthews and D. H. Saklofske (London: Sage Publications, Inc.), 261–292. doi: 10.4135/9781849200479.n13

Thayer, R. E. (1989). *The Biopsychology of Mood and Arousal*. (Oxford: Oxford University Press).

Võ, M. L. H., Conrad, M., Kuchinke, L., Urton, K., Hofmann, M. J., and Jacobs, A. M. (2009). The Berlin affective word list reloaded (BAWL-R). *Behav. Res. Methods* 41, 534–538. doi: 10.3758/BRM.41.2.534

Warriner, A. B., Kuperman, V., and Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behav. Res. Methods* 45, 1191–1207. doi: 10.3758/s13428-012-0314-x

Watson, D., and Clark, L. A. (1984). Negative affectivity: the disposition to experience aversive emotional states. *Psychol. Bull.* 96, 465–490. doi: 10.1037/0033-2909.96.3.465

Watson, D., Clark, L. A., and Carey, G. (1988). Positive and negative affectivity and their relation to anxiety and depressive disorders. *J. Abnormal Psychol.* 97, 346–353. doi: 10.1037/0021-843X.97.3.346

Watson, D., and Tellegen, A. (1985). Toward a consensual structure of mood. *Psychol. Bull.* 98, 219–235. doi: 10.1037/0033-2909.98.2.219

Wierzbicka, A. (1994). “Emotion, language, and cultural scripts,” in *Emotion and Culture: Empirical Studies of Mutual Influence,* eds S. Kitayama and H. R. Markus (American Psychological Association), 133–196.

Wierzbicka, A. (1999). *Emotions Across Languages and Cultures: Diversity and Universals*. Cambridge: Cambridge University Press.

Wojciszke, B. (1997). Parallels between competence- versus morality-related traits and individualistic versus collectivistic values. *Eur. J. Soc. Psychol.* 27, 245–256. doi: 10.1002/(SICI)1099-0992(199705)27:3<245::AID-EJSP819>3.0.CO;2-H

Wojciszke, B. (2005). Morality and competence in person- and self-perception. *Eur. Rev. Soc. Psychol.* 16, 155–188. doi: 10.1080/10463280500229619

Wojciszke, B., Dowhyluk, M., and Jaworski, M. (1998). Moral competence-related traits: how do they differ? *Polish Psychol. Bull.* 29, 283–294.

Wojciszke, B., and Szymków, A. (2003). Emotions related to others’ competence and morality. *Polish Psychol. Bull.* 34, 135–142.

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.

frontiers in Psychology

ORIGINAL RESEARCH
published: 20 October 2021
doi: 10.3389/fpsyg.2021.737756

Check for updates

Micropoetry Meets Neurocognitive Poetics: Influence of Associations on the Reception of Poetry

Katharina Gloria Hugentobler and Jana Lüdtke*

Department of Education and Psychology, Experimental and Neurocognitive Psychology, Freie Universität Berlin, Berlin, Germany

Reading and understanding poetic texts is often described as an interactive process influenced by the words and phrases building the poems and all associations and images induced by them in the readers mind. Iser, for example, described the understanding process as the closing of a good Gestalt promoted by mental images. Here, we investigate the effect that semantic cohesion, that is the internal connection of a list words, has on understanding and appreciation of poetic texts. To do this, word lists are presented as modern micropoems to the participants and the (ease of) extraction of underlying concepts as well as the affective and aesthetic responses are implicitly and explicitly measured. We found that a unifying concept is found more easily and unifying concepts vary significantly less between participants when the words composing a micropoem are semantically related. Moreover these items are liked better and are understood more easily. Our study shows evidence for the assumed relationship between building spontaneous associations, forming mental imagery, and understanding and appreciation of poetic texts. In addition, we introduced a new method well-suited to manipulate backgrounding features independently of foregrounding features which allows to disentangle the effects of both on poetry reception.

Keywords: neurocognitive poetics, associations, literary reading, text comprehension, computational linguistics

OPEN ACCESS

Edited by:
Francesca M. M. Citron,
Lancaster University, United Kingdom

Reviewed by:
Alexander Onysko,
University of Klagenfurt, Austria
Peter Dixon,
University of Alberta, Canada

****Correspondence:***
Katharina Gloria Hugentobler
hugentobler@zedat.fu-berlin.de

Specialty section:
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

Received: *07 July 2021*
Accepted: *14 September 2021*
Published: *20 October 2021*

Citation:
Hugentobler KG and Lüdtke J (2021) Micropoetry Meets Neurocognitive Poetics: Influence of Associations on the Reception of Poetry. Front. Psychol. 12:737756. doi: 10.3389/fpsyg.2021.737756

1. INTRODUCTION

> *The text, we have seen, patterns and delimits, but it ultimately functions like a chemical element: it itself is merged in the synthesis with the other elements to produce a particular event-a poem.*
> – Louise M. Rosenblatt

We, as a universal society, have long agreed that a poem is more than a mere collection of words: it is, we might feel, a picture painted with words. Psycholinguistic and neurocognitive research has only just started to explore the underlying cognitive and affective processes which help us make sense of a poetic text (Jacobs and Kinder, 2015; Jacobs and Willems, 2018) and ultimately shape and form mental images often reported while reading poetry (Belfi et al., 2018; Papp-Zipernovszky et al., 2021).

As already assumed by proponents of the readers-response theories like Iser (1978) and Rosenblatt (1978), reading and especially reading literature can only be understood as an interaction between features of the text and characteristics of the reader, and, as recently highlighted, also by the task at hand and the reading situation (Jacobs, 2015a; Bohn-Gettler, 2019). As described in the Transactional Theory by Rosenblatt (1978), reading of a poetic text is a process of meaning construction in which the linguistic symbols, the words and the written or printed marks have to be distinguished from what a listener or reader evokes from them (Rosenblatt, 1978). Rosenblatt used the term poem only for the later one. Iser described the meaning construction process as "one of the activities through which we form the 'gestalt' of a literary text" based on "the 'picturing' that is done by our imagination" (Iser, 1972, p. 288). But how does this interaction between reader and text come about? Let us start our analysis by looking at the interaction between text and reader from afar: At low resolution the emphasis is on the influence of literature genre on reading behaviour. According to proponents of the Reader Response Theory, reading behaviour depends on the reader's assumptions about the text (Hanauer, 1998; Topping, 2015) and a reading mode relevant to the understanding of the genre is adopted according to the genre choice made. As choices and adaptation of reading mode take time, several empirical studies demonstrated that texts labelled as factual are read faster compared to fictional texts (Zwaan, 1994; Wolfe, 2005; Altmann et al., 2014, see Hartung et al., 2017, for contrasting results). Also memory for individual aspects of a text differs as a function of the type of text (Zwaan, 1994). In line with the general assumption that classification of a text as a poem gives rise to a reading mode characterised by heightened awareness of language patterns and hidden meanings (for a comprehensive overview see Hanauer, 1998 and Carminati et al., 2006). Osowiecka and Kolanczyk (2018) found that poetry reading stimulates divergent thinking and vice versa.

According to Graesser (2015), the above theories are sufficient to explain basic levels of comprehension, but they are insufficient to describe deeper comprehension processes assumed to play a role when reading literature. The latter requires to also consider parameters such as the influence of emotions, situations etc. both through the literal content in itself, but also through an interaction between reader and text. In other words, text-bound variables like rhythm (Kentner, 2012; Peelle and Davis, 2012), prosody (Paulmann et al., 2013), valence and arousal within the words employed (Kuchinke et al., 2005; Warriner et al., 2013; Jacobs et al., 2016; Sylvester et al., 2016) and reader-centered variables such as affective and aesthetic responses (De Jaegher and Di Paolo, 2007; Gallagher, 2012; Winkielman et al., 2014), often studied from an embodied perspective, are closely intertwined. The Neurocognitive Poetics Model (NCPM; Jacobs, 2011, 2015a,c; Jacobs and Kinder, 2017) aims at integrating data on neurocognitive and emotional (Jakobson, 1960; Panksepp, 2008) processes during literature comprehension. It is conceptualised as a dual-process model which is informed by Gestalt theory and as such distinguishes between back- and foreground elements. The understanding of a text is conceived as successful closing of a good Gestalt. Within the framework of the model a text consists of familiarising background elements which cater to the readers need for coherence and meaning construction. The model further assumes that backgrounding features are stronger associated with immersive reactions (Van Krieken et al., 2015) and a mood-empathic reading mode (Lüdtke et al., 2014; Jacobs et al., 2016). Emotional and aesthetic feelings (Fitch, 2009; Westphal-Fitch and Fitch, 2018) are created against this background through foregrounding text elements. Foregrounding elements have a defamiliarising effect (Kuiken et al., 2012; Jacobs, 2015a,b,c) thereby eliciting what Mukarovský (1964) described as a de-automatisation in the reading process, leading to reduced reading speed as a result of more and intensive active drawing of inferences. Although manipulation of words always influences both backgrounding and foregrounding elements, empirical studies on reading poetic texts primarily focussed on the manipulation of foregrounding elements. Hakemulder (2004) for example changed isolated lines of poems to manipulate the amount of foreground features and observed that foregrounding caused higher aesthetic appreciation. Obermeier et al. (2013) manipulated rhyme and metre and also demonstrated that both foregrounding features enhanced aesthetic appreciation. As a text is characterised by both background and foreground features, experimental studies on the influence of background features like word frequency are equally important for the study of poetry comprehension. Yet studies focussing on the isolated manipulation of background features are few and far between (Hanauer, 1998). This raises the question of whether background features have any influence on comprehension of poetic texts at all? The short answer is: they do. For the long answer we need to look at studies addressing these elements.

Facilitation or ease of processing, i.e., the ease of mental operations concerned with stimulus meaning and its relation to semantic knowledge structures, is known to affect perceived aesthetic value. This is recognised as the mere-exposure effect, which links repetition of a stimulus to enhanced liking for this initially neutral stimulus (Zajonc, 2000). In support of this idea, Menninghaus et al. (2017) showed that parallelistic diction, a supra-lexical background feature, enhances emotional and aesthetic appreciation as well as semantic processing. In a similar way Reber and Schwarz linked variables that facilitate the processing of a stimulus (e.g., priming, mastery and the amount of information) to positive affective reactions (Reber et al., 2002, 2004). Markedly, aesthetic appreciation, a bona fide reaction to the presence of foregrounding elements, is reciprocally influenced by and at the same time acts as a positive influence on processing fluency (Reber et al., 2004). Moreover, the absence of prototypical narrative features (which are often discussed as an example of backgrounding features Hakemulder, 2020) strongly increases processing efforts as measured by reading times (Castiglione, 2017). Manipulation of word frequency, perhaps the most prototypical backgrounding feature, is widely employed to highlight the influence of facilitating factors on processing speed and accuracy (Murray and Forster, 2004) and appears to mediate pleasentness- and familiarity ratings on single word level (Sluckin et al., 1980).

Although ease of processing was shown to influence the above, the concept is not sufficient to explain the closing of a good Gestalt *sensu* (Iser, 1972). While an explanation based on the ease of processing focusses on single word level cognition (e.g., higher word frequency leads to faster word recognition), a good Gestalt results from the integration of words. In the same way that cognitive theories assume that associations are the most important principle for organisation of our semantic knowledge (for an overview cf. Hofmann and Jacobs, 2014), they should also guide the closing of a good Gestalt while reading poetic texts. In this regard, we can posit an influence of what James (cited after Mangan, 2001, 2008) described as the fringe of consciousness on processing fluency-mediated construction of meaning: It represents context information in consciousness, conveys a sense of rightness and brings out the notion that more information is available.

What do we already know about the role of associations on text comprehension processes related to the closing of a good Gestalt? According to the Gestalt principle of closure, objects which are grouped together are perceived as a whole which might otherwise (i.e., if the objects were regarded as individual entities) not exist and which is more than or different from the combination of its parts. The emergence of a resulting Gestalt from an albeit sketchy figure is hinged on the presence of a unifying pattern (Warriner, 1923). Wolfgang Iser transmitted that principle onto the process of reading literary texts and therefore poetry: "By grouping together the written parts of the text, we enable them to interact, we observe the direction in which they are leading us, and we project onto them the consistency which we, as readers, require. This 'gestalt' must inevitably be colored by our own characteristic selection process" (Iser, 1972, p. 289). A poem, composed of words that interrelate, facilitates the application of the grouping principle of closure so that a good Gestalt can emerge and potentially foster the creation of a mental image (Iser, 1972).

Yet, words have an effect on comprehension in their own right. Readers pre-activate concepts beyond the unfolding sentence to simulate the described event for predictive (online) language comprehension (for an overview see Huettig, 2015). Associations or mental connections between concepts or events are by some described as the product of contiguity, i.e., the probability of co-occurrence between words. Others conceptualise associations as the result of an event simulation heuristic which pre-activates linguistic representations (Huettig, 2015). Semantic cohesion serves as a measure for the strength of internal association between concepts and was shown to influence sentence processing on single sentence level (Hintz et al., 2020). Current machine-learning methodologies (Grainger and Jacobs, 1996; Dijkstra and van Heuven, 2002; Hofmann and Jacobs, 2014; Jacobs, 2018) as well as some association databases rely on contiguity-based approaches. Accordingly, in a recent study Jacobs and Kinder showed that latent semantic analysis (LSA, a method based on distributional semantics to predict upcoming words) is superior to surface model analysis for the prediction of ratings on difficulty and aesthetic liking of metaphors (Jacobs and Kinder, 2017). This reveals the power of associative databases which exceed purely systematic relations

result	transplant	bed
banana	beat	dream
fierce	love	awake
active	soul	night
tape	ache	rest
cigarette	throb	peace
equality	attack	sound
poor	pump	tired
charm	red	death

FIGURE 1 | Three examples of items presented as micropoems in this study. Items are ordered according to their semantic cohesion (in parenthesis), from left to right: Match 4 (0.208), EAT 3 (0.264), EAT 5 (0.369). For a comprehensive overview over the items please cf. **Table 1**; **Supplementary Tables 1, 2**.

(such as WordNet) and instead rely on a rich network of semantic relations informed by cultural, emotional and personal experience (Netzer et al., 2009).

Poems lend themselves to the study of neurocognitive processes during reading, seen as they are well constrained literary material (Müller et al., 2017). Seen in the light of associative networks, poetry appears to be characterised by lexical associations (Netzer et al., 2009). Correspondingly, a computational analysis of poems written by professional and amateur poets revealed that the former use more concrete words to create a canvas onto which readers can project their associations. Professional poets avoid direct references to abstract and intangible concepts and instead let images instead of words convey emotions, concepts, and experiences (Kao and Jurafsky, 2012).

With the above in mind, we want to investigate the effect of associations, as an example of a backgrounding feature, on the closing of a good Gestalt and maintain that the effect exceeds a mere effect of facilitated processing. To test this proposition, we created simple word lists, which are essentially devoid of foreground features, and presented them as versions of poetic texts (**Figure 1**). The participants are asked to read lists of nine words with varying degree of semantic cohesion introduced as micropoems (Jacobs, 2015b). They are then asked to produce a possible one-word-title for each poem and rate the micropoems on dimensions Comprehensibility, Liking, Imageability as indication of reader's response at both the upper and lower route according to the NCPM (Jacobs, 2015a). This setup allows for an analysis of the influence of background features -more precisely the role of semantic cohesion and word frequency in the present case- on the closing of a good Gestalt independent of the influence of foregrounding figures.

2. HYPOTHESES

As each and every word is linked to a plethora of other concepts, reading such a simple word list not only evokes simulation of the mentioned concepts but also co-activates the concepts associated with those words (Mangan, 2001; Hofmann and Jacobs, 2014). Each group of concepts associated with a single word in itself is likely to evoke its own association, seen as they are all semantically related and trace back to a common anchor word (Deese, 1959; Roediger and McDermott, 1995). The idea behind our set-up is *not* that each word *per se* leads to the generation of many associations. Instead, the words taken together are all highly semantically related (high semantic cohesion as a feature facilitating formation of a Gestalt) and thus, as a whole, facilitate the generation of associations in the reader's mind. We therefore assume a reciprocal enhancement of activation which facilitates the formation of a consistent mental image or good Gestalt as for example opposed to a serial position effect. If only the first or the last word of a poem was used for the generation of a title, we should not observe any effect of semantic cohesion on semantic relatedness. Reading poems with low internal semantic cohesion, on the other hand should merely activate concepts vaguely associated with the individual words. What is more, associations may neutralise each other through lateral inhibition at the conceptual level (for a thorough analysis see Hofmann and Jacobs, 2014) and therefore inevitably fail to close a good Gestalt.

Consequently, we should be able to test whether semantic cohesion of the words composing a poem has a significant impact on the closing of a good Gestalt, and therefore of the formation of mental imagery, a process associated with the recognition of the bigger picture conveyed by a poetic text. Hence, we expect participants to understand the micropoems with higher semantic cohesion more easily and additionally report the experience of more mental imagery. Moreover, semantic cohesion should also influence affective and aesthetic responses as described in the NCPM (Jacobs et al., 2015). Following the scheme introduced by Dixon and Bortolussi (2015) a set of explicitly surveyed rating data (mood induced, liking of the micropoem, perceived difficulty of the micropoem, perceived imagery induced) and ancillary reaction parameters (mean semantic relatedness of the response, mean reading time) as implicit control variables will be used.

In accordance with the assumptions of the NCPM, a modification of micropoems in terms of internal semantic cohesion should be reflected in the mood induced in the reader (NCPM: backgrounding, fast route) and self-reported aesthetic appreciation of the micropoems (NCPM: foregrounding, slow route) (Jacobs, 2015a). With regards to difficulty, the effect should be mirrored in the self-reported difficulty to understand the poems (explicit measure) and reading times (implicit measure). The capacity of a poem to induce a mental image in the reader should be seen in a self-reported rating of perceived mental images. Moreover, participants will be asked to put their first spontaneous association to the poems forward in form of a freely chosen title. Semantic relatedness between spontaneous associations to the poems (i.e., titles) and the words composing a poem as well as the number of different titles found, as an indicator for the closing of a good Gestalt, are used to control this effect with an implicit measure.

One might argue that the effects discussed above can be put down to facilitated processing due to semantic cohesion. Following this line of thought, if the root cause of any differences observed was indeed attributable to an ease of processing, this effect should reproduce if another factor influencing the ease of processing (e.g., supra-lexical word frequency) was modified. If supra-lexical word frequency has no or a different effect, then it is clear that the influence of semantic cohesion is based on more than a simple facilitation of word identification.

We expect self-reported mood and imagery to be stronger for micropoems with high semantic cohesion than low semantic cohesion. Micropoems with high semantic cohesion should be easier to understand (less self reported difficulty in understanding) and should be better liked. Regarding the implicit measures we expect higher semantic relatedness between the title and the reported associations (i.e., titles given) and faster reading times for micropoems with high internal semantic cohesion. Moreover the intra-subject variance of associations found should increase for micropoems with reduced internal semantic cohesion.

We attribute the effects to the evolution of more elaborate associations in the participants' minds as a function of internal semantic cohesion and therefore do not expect any effect for word frequency.

3. METHODS

3.1. Participants

The sample was composed of 32 participants (20 female; age: $M = 46.13$, $SD = 13.96$, range = 24–74) all of whom learned English as their first language. Participants were recruited through an online survey employing the www.soscisurvey.de platform. A total of 9 participants reported to have been exposed to a language other than English before the age of 5, while 27 of 32 participants lived in an English speaking environment at the time of the survey and 28 participants used English as their most frequent language. The experiment followed the rules set by the ethical guidelines of the German Psychological Society's 121 (DGPs, 2004, CIII). Participants were informed about taking part in research, about the possibility of quitting the experiment with no disadvantage at any time and about the fact that all data were collected and analysed anonymously. They provided informed consent and gave permission to use their collected data anonymously for publications.

3.1.1. Inclusion Criteria

The sample was stratified by first language (BE/AE/undefined), second language (yes/no), current immersion in English speaking environment (yes/no) and English as dominant language (yes/no). Participants who spoke any language other than English as their first language were excluded by default ($n = 6$). Data quality was checked based on the mean semantic distances obtained or the titles produced by the participants and the words constituting a poem (for calculation, please see section 3.4.2). If a participant had entered titles at random the corresponding mean semantic distance of their answers should differ from the global mean. We did not observe such behaviour in the data, therefore all remaining $n = 32$ (after exclusion of $n = 6$ non-native speakers) participants were included into the final dataset. This constitutes a 84.21% of the original $n = 38$ sample.

3.2. Item Design

We created a first set of micropoems with high semantic similarity, which consisted of nine words each. The micropoems were drawn as subsamples of the Edinburgh Associative Thesaurus, a database for 8,400 stimulus words together with their most frequent word associations, which were collected in an empirical free-association set-up based on 100 subjects (Kiss et al., 1973).

To select a set of six micropoems the following criteria were used. We first selected only those stimulus words from the original EAT for which nine different word associations (maximum number of associations reported) were listed. We further reduced the list of stimulus words by excluding all stimulus words with missing values for word frequency (taken from the SUBTLEX-UK, van Heuven et al., 2014) and word-based valence, arousal, concreteness and imagery (taken from Bestgen and Vincze, 2012) for the nine word associations. For the resulting list of stimulus words we calculated the supra-lexical values for the associated micropoems by calculating the mean values for word-based valence, arousal, concreteness and imagery for the nine word associations constituting each micropoem. We then excluded all stimulus words and associated micropoems with supra-lexical valence, arousal and concreteness values outside the global interquartile range (IQR) of the dimensions valence [4.2, 6.1] and arousal [4.82, 5.36] calculated for all words reported by Bestgen and Vincze (2012).

To manipulate the second predictor, the supra-lexical level of word frequency, we looked for micropoems with high and low supra-lexical word frequency. Although the correlation between word frequency and word imageability is in general very low (Stadthagen-Gonzalez and Davis, 2006), we explicitly tried to avoid any confoundation between word frequency and imageability. Consequently, we first used the supra-lexical imagery for the associated micropoems of the remaining stimulus words and divided them in three groups: a low imagery, a medium imagery and a high imagery one. For doing so we used the first quartile, the IQR, and third quartile for word imagery of the Bestgen and Vincze list as defining criteria. In a final step we chose poems with a supra-lexical word frequency above, within, and below the IQR of lexical word frequency taken from SUBTLEX-UK [2092; 33275].

We controlled the respective numerical semantic cohesion (cf. section 3.3 below) and found that although the micropoems were composed of highly associated words (determined in a free association set-up), they show substantial differences between empirical semantic association and numerical semantic cohesion. Therefore, in a second step we constructed six new micropoems constituted of a list of nine words with no or even low semantic associations and controlled for semantic cohesion as well as independence of semantic cohesion and word frequency. To make sure, that the six EAT-based micropoems (EAT-1 to EAT-6) and the six matching micropoems (Match-1 to Match-6) do not differ with respect to the supra-lexical values of interest, we used the program Match (Van Casteren and Davis, 2006) to randomly pick nine words matching the nine words constituting the EAT-based micropoems on dimensions valence, arousal, imagery, concreteness, frequency, and word length. The drawing pool for the words constituting the Match-based micropoems consisted of all stimulus words of the EAT for which all psycholinguistic values mentioned above were available. Words already part of one of the six EAT-based micropoems and words, that were both noun and verb (e.g., "fight"), were excluded.[1] The resulting twelve EAT-based and Match-based micropoems as well as their relevant measures at the lexical level are reported in **Supplementary Tables 1, 2**. The supra-lexical values for all micropoems are summarised in **Table 1**.

3.3. Measures for Semantic Similarities

We refer to the dependent variable as "semantic relatedness" in order to distinguish it from the independent variable "semantic cohesion." In general, semantic similarity between two words is calculated based on a 300 dimensional vector space model trained with the fastText-based skipgram algorithm (Bojanowski et al., 2017) and based on the Gutenberg English Poetry Corpus (GEPC; Jacobs, 2018) using the similarity function of the gensim library (Řehůřek and Sojka, 2010) in Python 3.7.

3.3.1. Determination of Semantic Relatedness

Semantic relatedness refers to our measure for the similarity between a title entered by the participants and all words of the corresponding micropoem. For each micropoem mean semantic similarity was calculated as individual mean of all semantic similarity values thus obtained. "Maths," an element of the micropoem "proof" was not part of the GEPC corpus (nor were mathematics and all corresponding spellings). Therefore mean semantic relatedness for micropoem EAT-4 was calculated based on the remaining eight other words of the micropoem EAT-4.

3.3.2. Determination of Semantic Cohesion

Semantic cohesion refers to the predictor variable and is assigned to the micropoems based on its construction only. It is determined as grand mean of semantic similarity between each word of a micropoem with all remaining other words of this micropoem (for individual semantic cohesion of the micropoems please see **Table 1**).

3.4. Data Collection

We introduced 12 items as examples of modern poetry to the participants. Because of the shortness, micropoems were presented as a vertically arranged list of words with one stanza (cf. **Figure 1**) and were displayed in random order. Participants were asked to give each micropoem a one-word title. Participants were explicitly instructed to avoid giving an element of the micropoem itself as a title. All titles given are summarised in **Supplementary Tables 3, 4**.

3.4.1. Predictors

Besides the manipulation of semantic cohesion ($M = 0.27$, $SE = 0.01$, range = 0.21–0.31), we used the supra-lexical word frequency, the mean word frequency of all words constituting a micropoem as a second continuous predictor ($M = 36294.42$, $SE = 6769.46$, range = 7444–82912). The individual values for both predictors for all micropoems are reported in **Table 1**.

[1] We realised that in one case the EAT-based micropoem EAT-3 and its corresponding Match-based micropoem Match-3 both contained one identical word ("love").

TABLE 1 | Supra-lexical word length, word frequencies, psycholinguistic rating values as well as semantic cohesion and number of different titles obtained for EAT-derived (EAT-1 to EAT-6) and matched (Match-1 to Match-6) items.

Item	Length	Frequency[a]	Valence[b]	Arousal[b]	Imageability[b]	Cohesion[c]	Ntitle[d]
Match-4	6	7,444	5.3	5.2	4.5	0.208	29
Match-1	5	33,772	5.7	5	5.1	0.225	27
Match-6	6.3	37,268	5.5	5.1	4.7	0.233	31
Match-5	4.6	27,688	5.1	5.1	4.9	0.239	26
Match-3	5	29,016	4.9	5.3	5	0.249	29
EAT-3	4.9	28,206	4.9	5.2	5.1	0.264	18
EAT-4	6.2	8,166	5.4	5.1	4.4	0.266	25
Match-2	5.7	80,781	5.8	5.2	4.8	0.266	31
EAT-1	5	33,902	5.7	5	5.1	0.289	20
EAT-2	5.6	82,912	5.7	5	4.8	0.296	26
EAT-6	6.3	38,108	5.5	5	4.6	0.337	20
EAT-5	4.7	28,270	5	5	5.1	0.369	16

[a] *Word frequencies were taken from van Heuven et al. (2014).*
[b] *Psycholinguistic rating values were taken from Bestgen and Vincze (2012).*
[c] *Intra-item semantic cohesion was determined as the grand mean of semantic similarity between each word of an item with all remaining other words of this item.*
[d] *Number of different titles found for each item.*

3.4.2. Implicit Measures

Reaction times for first time reading as well as mean semantic relatedness between each title given and the corresponding words of each micropoem were collected as implicit measures for understandability of the micropoems presented. Reaction times were collected for first time reading of each micropoem. The time between loading of the page and the moment when the participant moved on to the next page was recorded. The precision of the measurement was 1 s. The reading time recorded for participant 143 and micropoem EAT-3 was 851 s. This is indicative of a temporary deflection of the participant's attention rather than a correct representation of the time taken to read the micropoem. Therefore the reading time for this particular micropoem and participant was approximated by the mean reading time of participant 143 for all other micropoems with high semantic relatedness.

3.4.3. Explicit Measures

Additionally, participants were asked to rate their own reading experience on 5-point rating scales for dimensions Difficulty, Imagery (Kuiken et al., 2012), as well as Liking of the micropoem and experience of Mood (Lüdtke et al., 2014, cf. **Table 2**). This was used as an explicit measure for participants' reception of the poems. The order of the questions was randomised for each participant. Additionally, reading habits and individual preference for poetry (modern vs. classic) as well as enjoyment of poetry were assessed for each participant.

3.4.4. Treatment of Invalid Titles

Titles for which the mean semantic relatedness value was not determined (either because the title entered was not part of the vocabulary or because it was a two-word title) were excluded in the linear mixed model analysis of semantic relatedness. This represents a 10.1% of the final dataset of all titles entered. All remaining titles were corrected for British orthography.

3.5. Statistical Analysis

A Box-Cox analysis using the MASS::boxcox() command (Venables and Ripley, 2002) incrementing by 0.1 within the interval [−6, 6] revealed a deviation from normal distribution of all dependent variables. The data was therefore transformed using the corresponding lambda (cf. **Supplementary Table 5**) and were used for inferential statistical analysis. Statistical analyses were conducted using linear mixed effects regression models (Baayen et al., 2008; Barr, 2013; Barr et al., 2013; Magezi, 2015) using the package lme4 (Bates et al., 2015) in the statistical environment R (R Core Team, 2020). Fixed effects of semantic cohesion and supra-lexical word frequency on dependent variables semantic relatedness of the titles, reading time, perceived liking, perceived mood, imagery experienced, perceived difficulty, and number of different titles found, respectively, were checked with Wald F-tests and a Kenward-Roger approximation of degrees of freedom using the car::Anova() function (Fox and Weisberg, 2019). Random effects were assumed for subjects and micropoem sets and were approximated allowing for individual intercepts for participants, and inclusion of random slopes where appropriate. In line with recent recommendations (Barr, 2013) complex models were generated for all variables (random slopes models and/or interaction models, where appropriate) and complex models were compared to models with sequentially decreased complexity to find the model fitting the data best. Model comparison was performed using the stats::anova() function. Unless otherwise indicated, the most parsimonious model, a linear mixed model with individual intercepts for random person and micropoem factors, was used for the final statistical analysis. A general linear model was used to predict the of number of different titles by semantic cohesion. This was calculated using the stats::lm() function. For all models, Nakagawa's R^2 was calculated using the performance::r2_nakagawa function (Nakagawa and Schielzeth, 2013). All data were analysed on a

TABLE 2 | Dimensions assessed and corresponding items in the poem questionnaire.

Dimension	Item	References
Liking	I like the poem.	Lüdtke et al., 2014
Experience of Mood	Reading the poem let me experience a mood.	Lüdtke et al., 2014
Imagery	While reading this poem the images that came to mind seemed pregnant with meaning.	Kuiken et al., 2012
Difficulty	I had to think long and hard about the poem before I could come up with a title.	This work

Items were assessed on a 5-point rating scale.

TABLE 3 | Correlation matrix of the dependent variables.

Variable	1	2	3	4	5	6	7
1. Mean semantic relatedness	-						
2. Reading times	−0.594	-					
3. Number of titles	−0.741	0.602	-				
4. Experience of Mood	0.545	−0.732	-0.507	-			
5. Liking	0.758	−0.788	−0.606	0.752	-		
6. Imagery	0.576	−0.866	−0.67	0.825	0.822	-	
7. Difficulty	−0.769	0.858	0.772	−0.629	−0.808	−0.869	-

TABLE 4 | Summary of the statistical analysis.

Variable	df_1, df_2	F	p	REM[a]	R^2 (conditional, marginal)
Mean semantic relatedness	1, 9.92	27.6	0.0004	Intercepts	0.413, 0.248
Mean reading time	1, 10.11	8.09	0.0172	Intercepts	–
Number of titles[b, c]	1, 10	12.2	0.0058	n.a.	0.55, 0.504
Experience of mood	1, 10	4.45	0.06	Intercepts	–
Liking	1, 10	13.54	0.004	Intercepts	0.462, 0.039
Imagery	1, 10	5.04	0.048	Intercepts	–
Difficulty	1, 10	13.3	0.0045	Intercepts	0.401, 0.042

Data were analysed on a Bonferroni-corrected alpha level of 0.05/7 = 0.007. Explicit and implicit measures are separated by a single line.
[a]*Random effects modelling: intercepts = individual intercepts.*
[b]*A general linear model was employed and semantic cohesion was used as a predictor for the number of different titles found.*
[c]*R^2 is given as multiple, adjusted.*

Bonferroni-corrected alpha-level of 0.05/7 = 0.007. Please refer to **Table 3** for a summary of the correlation matrix for the different dependent variables and to **Table 4** for a summary of all final models.

4. RESULTS

We have always started our analysis with a complex model with maximal degrees of freedom for the random effects. All initial models contained both fixed effects, i.e., semantic cohesion and supra-lexical word frequency and their interaction (cf. **Table 1**). We then reduced the random effects model in stepwise manner until the model was no longer overfitting (Baayen et al., 2008; Barr, 2013; Barr et al., 2013). For all dependent variables interaction or main effects with supra-lexical word frequency never became significant (all $F < 0.4$, all $p > 0.5$) and were therefore removed.

4.1. Objective Measures

A positive relationship between semantic cohesion and semantic relatedness as a proxy for the closing of a good Gestalt (cf. **Figure 2A**) revealed to be significant (cf. **Table 4**, entry 1). A slightly negative relationship was observed between mean reading time and semantic cohesion (cf. **Figure 2B**) but did not become significant on a Bonferroni-corrected alpha level (cf. **Table 4**, entry 2). A strongly negative relationship between semantic cohesion and number of different titles (cf. **Figure 2C**) revealed to be significant on the Bonferroni-corrected alpha-level (cf. **Table 4**, entry 3).

4.2. Subjective Measures

A moderately positive relationship between semantic cohesion and participants' reported perception of mood (cf. **Figure 3A**) did not reveal to be significant on the Bonferroni-corrected alpha-level (cf. **Table 4**, entry 4). A moderately positive relationship was found between semantic cohesion and participants' reported liking of the items (cf. **Figure 3B**) and was significant on the Bonferroni-corrected alpha-level (cf. **Table 4**, entry 5). Moreover, the moderately positive relationship between semantic cohesion and participants' reported perception of mental images (cf. **Figure 3C**) did not reveal to be significant on the Bonferroni-corrected alpha-level (cf. **Table 4**, entry 6), while the negative relationship between semantic cohesion and participants' reported perception difficulty of understanding the items (cf. **Figure 3D**) was significant on the Bonferroni-corrected alpha-level (cf. **Table 4**, entry 7).

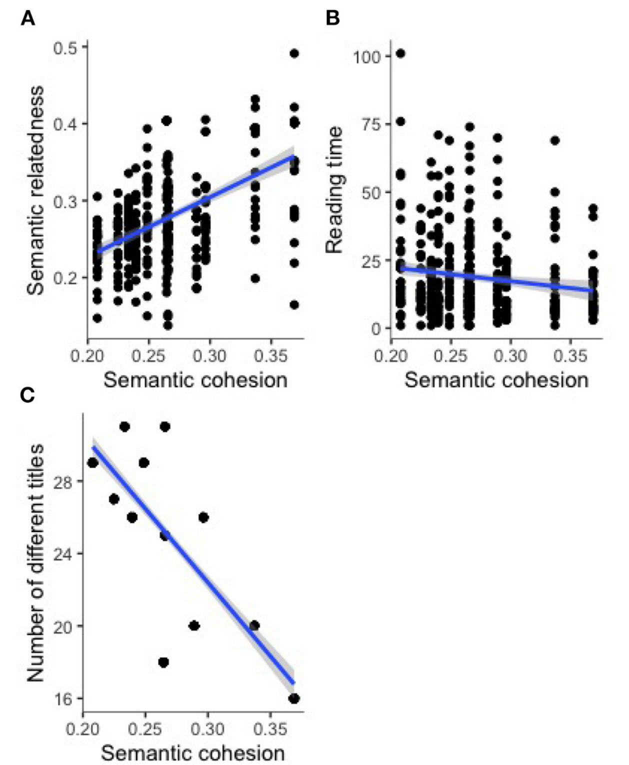

FIGURE 2 | Explicit measures: Relationship between semantic cohesion and semantic relatedness between titles and poems **(A)**, mean reading time **(B)**, and number of different titles **(C)**.

5. DISCUSSION

We have set out to understand the mechanisms that make us, as readers of a poem, see a bigger picture upon the inward eye. We hypothesised that even background features should have a decisive influence on reading and understanding poetic texts. More precisely, we assumed that semantic relatedness of the words making up a poem, as one example for a supra-lexical background feature, is a main factor governing the formation of the closed Gestalt. To test this hypothesis we had participants read lists of words with different amount of semantic relatedness and assayed their emotional response as well as their associations as a proxy for the mental images produced. Moreover we also manipulated supra-lexical word frequency defined as the mean word frequency of all words composing the micro poem, to test whether possible effects rely on a simple facilitation of word identification only. Taken together, we found no effect for supra-lexical word frequency but an effect for semantic cohesion on all objective variables as well as on subjective ratings liking and difficulty.

The observed differences of semantic relatedness between titles and micro-poems as a function of semantic cohesion lets us understand two things: first we can safely conclude that our manipulation was successful. Secondly, on condition that semantic relatedness of the titles is indicative of the capacity of a poem to induce a mental image, we have shown that the closing of a good Gestalt is at least partially hinged on semantic cohesion of a micropoem. Up to this point a mere effect of facilitated processing would be an equally plausible explanation for the present results. If that was the case, however, supra-lexical word frequency, which famously facilitates processing

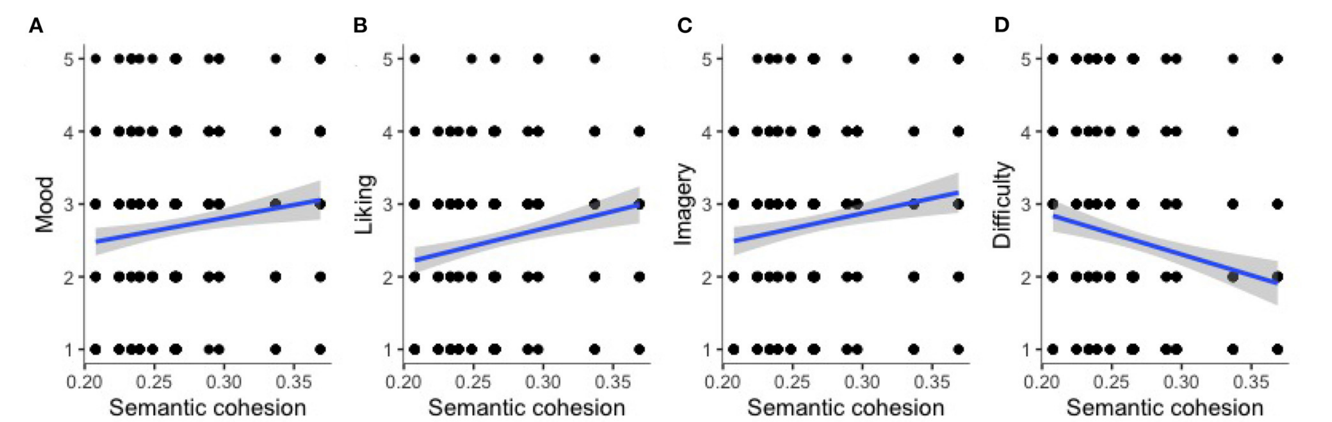

FIGURE 3 | Implicit measures: Relationship between semantic cohesion and participants rating of perceived mood **(A)**, liking of the items **(B)**, imagery **(C)**, and perceived difficulty in understanding **(D)**.

of words and texts (Lüdtke et al.)[2] should lead to the same results. We, however, did not find any main or interaction effect of word frequency on any variable examined. Therefore the present results need to be explained through effects beyond facilitation of processing. Our findings are in line with the results of Jolsvai et al. (2013), showing that meaningfulness of sentence fragments is a superior predictor for language processing speed than word frequency. Taken together this indicates that regarding the formation of a bigger picture facilitation of lexical access, i.e., basic comprehension processes at simple word level, is a second player to integration processes. In the following we will discuss this in more detail in reflection of current discourse on mental imagery, aesthetic appreciation, and comprehension difficulty.

Let us begin with an analysis of the results concerning mental imagery. Further to semantic relatedness, we analysed the number of different titles between items with high and low semantic cohesion, as a direct measure of mental imagery and the closing of a good Gestalt. Items with high semantic cohesion were expected to point to a target word or induce a target concept and therefore to yield a lower number of different associations. Items with low semantic cohesion on the other hand should lead to the generation of more different and less related associations. The present data confirmed this hypothesis: items with high semantic cohesion yielded fewer and less semantically related associations. While the analysis of the number of titles would indicate that a good Gestalt was indeed derived from items with high semantic cohesion, directly assayed generation of mental imagery (see above) did not. It can only serve as indicator or a trend pointing in the same direction. Although the latter finding at first glance seems to contradict our hypothesis, it is, upon closer inspection, a faithful reflection of the ongoing discussion whether mental imagery is indeed produced during reading, or is more of a literal afterthought (Jacobs and Willems, 2018). Moreover, the poems were essentially devoid of a deeper meaning or plot (Teng et al., 2016). This already by and in itself may disturb the process of closing a good Gestalt and thereby have interfered with the creation of mental images. In fulfillment of Rosenblatt's Transactional Theory (Rosenblatt, 1978), we also may not neglect the reader in this equation. Reader-based variables, particularly their response to the set task and the consequently adopted reading goal, can increase a reader's tendency to employ an inferential reading mode (Bohn-Gettler, 2019). In light of this, the task set here may have distracted the reader from adopting an inferential reading mode which otherwise would have allowed for mental images to be drawn. The readers may instead have focussed on *getting it right* and may therefore have lost access to any actively communicable notion of a mental image. Looking at participants' ratings for aesthetic appreciation of the items, no effects were found on dimension experience of Mood, but clear effects were obtained for Liking. This is in opposition to the predictions derived from the NCPM, according to which manipulation of background features should be reflected in upper route processes more than in lower route processes. Back- and foregrounding elements, however, cannot be thought as entities isolated from each other, seen as they mutually affect each other (Reber et al., 2004). In this sense, Obermeier et al. (2013) showed that manipulation of foreground features can influence upper route processes. Also Lüdtke et al. (2014) demonstrated cross-communication between foregrounding and backgrounding elements. These findings point in the direction already outlined by Jacobs (2011) suggesting that it is the quotient of backgrounding and foregrounding elements more than their individual contribution that facilitates enhanced processing along either route. More research is needed to shed light on this effect. In this sense, our approach lends itself to the analysis of the influence of semantic activation on both trajectories through the use of a set of micropoems with continuously increasing semantic cohesion. Moreover Belfi et al. (2018) showed that vividness of mental imagery has a crucial effect on aesthetic appreciation of poems.

On the axis Comprehensability, the subjective measure (survey data) became significant, while the objective measure (reading time) just failed to be significant on a Bonferroni corrected alpha level, but still indicated a correlation on a descriptive level. Reading time could, theoretically, be construed as a measure for ease of processing. By showing that word frequency, a factor directly related to processing speed, does not have any influence on reading times we demonstrate that semantic cohesion does not only influence basic language processing like word identification but presumably also higher order processes like integration and the closing of a good Gestalt. We argue that micropoems encouraging the formation of mental images should be easier understood and hence read faster. Magyari et al. (2020), per contra, reported prolonged fixation duration on words associated with an enactive style text, for which participants reported more mental images. On a fist glance this, again, seems to contradict our results which indicated faster reading times for items that were reported to induce more mental images. Yet, processing a prose text differs fundamentally from processing of our micropoems. Reading poetic texts is characterised by enhanced re-reading (Xue et al., 2020). Therefore the reading times measured in this study not only reflect ongoing reading processes (i.e., first pass reading and fixation duration as captured in the eyetracking data reported by Magyari et al., 2020) but also higher order comprehension processes necessary for fulfilling the task at hand. Contrary to studies on literary reading, in our case subjects had to find one word titles. We assumed that semantic cohesion fosters the closing of a good Gestalt which leads to enhanced mental imagery. This should simplify the task and thereby enhance participants' reaction times. More studies need to sheet light onto the role of mental imagery in reading prose and poetry. The method employed here is a promising tool to account for factors influencing mental imagery when reading poetic texts.

5.1. Limitations and Outlook

The Edinburgh Associative Thesaurus, which we used as database for the generation of our items with high semantic

[2]Lüdtke, J., Fröhlich, E., and Jacobs, A. M. (submitted). Small words matter: An eyetracking study about reading affective words in context. *Front Psychol.*

cohesion was created in the 1970ies and therefore may not fully reflect the current-day associative landscape. The use of recent computational tools such as fastText (Bojanowski et al., 2017) or GloVe (Pennington et al., 2014) in combination with a modern corpus is a preferable choice for the generation of items with high semantic cohesion. The current data show that semantic cohesion influences the generation of associations as well as liking (and possibly by extension also mood and mental imagery) through mechanisms which exceed mere facilitation of processing. Yet, our data show that high semantic cohesion does not automatically lead to the generation of mental images. As much as we expect emotions to be dependent on text, task, and reader (Rosenblatt, 1978; Bohn-Gettler, 2019) we may claim the same to be true for mental imagery. This should be considered in the item construction and design of future studies in order to disentangle the individual influences of person-, text-, and task-related factors on the formation of mental images. Moreover the perceived quality or meaningfulness of the items employed should be verified in a pre-study.

On a formal-methodological account, the method established in this work is well-suited to study the much-discussed question of when and where mental images are generated during the reading of poems. Our items were specifically designed to serve as a "minimal example" of a poem, which is devoid of all features of a traditional poem except one: semantic cohesion. This was deliberately done to minimise confounding effects through other variables traditionally present in poetry. The use of such items facilitated the isolated analysis of the effect of semantic relatedness on the closing of a good Gestalt. Moreover the items, devoid by design of foregrounding elements, gave us a tool to singly manipulate background features. Our results demonstrated that simple manipulation of background features is possible in a set-up with short stimuli. Our method is therefore relevant for fMRI studies, which can help to explore the assumed higher order processes during online-formation of mental imagery.

What is more, our method is particularly tailored to manipulate and target the complex interaction between background features and text processing along the upper and lower route postulated by the NCPM. This ultimately allows to conceptualise backgrounding as more than an instance from which literary devices called foregrounding elements stand out (Hakemulder, 2020).

Nevertheless, the effects observed in this study ought to be verified using longer items with more stylistic elements normally associated with poetic texts such as rhyme schemes, meter and verse form. Based on our findings we expect also higher appreciation and better understanding for more poem-like material with high semantic cohesion compared to low cohesion. We instructed our participants to treat our items as an example of modern poetry, which should have an influence on the reading mode, as for example, demonstrated by the fact-fiction paradigm (Altmann et al., 2014). Yet based on the current data we cannot rule out that some participants interpreted our items as arbitrary word lists instead of poems. Future studies using fMRI might answer that question: If the micropoems were processed as poetic texts, activation in the bilateral precentral and inferior frontal gyrus, as well as the right dorsolateral prefrontal cortex, anterior insula, and temporal pole should be observed (Jacobs and Willems, 2018).

5.2. Conclusion

Each and any of our analyses leads to the conclusion that associations, although they may have an influence on early processes, shape and guide higher integration processes. We therefore suggest that the bigger picture, as it were, of a poem is coloured by a collection of associations.

Psycholinguistic elements of a text amalgamate to create something larger which we, as readers, experience as *a poem*. We have shown that the synthesis of a poem from its elements exceeds the effect of facilitated processing. Significantly higher reported liking of items with high semantic cohesion and reduced perceived difficulty in understanding of the same items in combination with the respective reduced number of different titles strongly suggests that semantic cohesion influences the closing of a good Gestalt and as such may catalyse the generation of mental imagery. Although items with high semantic cohesion, devoid of meaning and plot, may not be sufficient to fully induce a mood and create communicable mental images, we have shown that they do pave the path toward the generation of mental images.

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 Ethik Kommission der Freien Universität Berlin. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

KH and JL contributed equally to the conception and design and report of the study. All authors have read and agreed to the final version of the manuscript.

FUNDING

We acknowledge support by the Open Access Publication Fund of the Freie Universität Berlin.

ACKNOWLEDGMENTS

The authors would like to thank Nicholas J. Weise for linguistic support.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2021.737756/full#supplementary-material

REFERENCES

Altmann, U., Bohrn, I. C., Lubrich, O., Menninghaus, W., and Jacobs, A. M. (2014). Fact vs fiction-how paratextual information shapes our reading processes. *Soc. Cogn. Affect. Neurosci*. 9, 22–29. doi: 10.1093/scan/nss098

Baayen, R. H., Davidson, D. J., and Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *J. Mem. Lang*. 59, 390–412. doi: 10.1016/j.jml.2007.12.005

Barr, D., Levy, R., Scheepers, C., and Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: keep it maximal Dale. *J. Mem. Lang*. 68, 1–43. doi: 10.1016/j.jml.2012.11.001

Barr, D. J. (2013). Random effects structure for testing interactions in linear mixed-effects models. *Front. Psychol*. 4:328. doi: 10.3389/fpsyg.2013.00328

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw*. 67, 1–48. doi: 10.18637/jss.v067.i01

Belfi, A. M., Vessel, E. A., and Starr, G. G. (2018). Individual ratings of vividness predict aesthetic appeal in poetry. *Psychol. Aesthet. Creat. Arts* 12, 341–350. doi: 10.1037/aca0000153

Bestgen, Y., and Vincze, N. (2012). Checking and bootstrapping lexical norms by means of word similarity indexes. *Behav. Res. Methods* 44, 998–1006. doi: 10.3758/s13428-012-0195-z

Bohn-Gettler, C. M. (2019). Getting a grip: the PET framework for studying how reader emotions influence comprehension. *Discourse Process*. 56, 386–401. doi: 10.1080/0163853X.2019.1611174

Bojanowski, P., Grave, E., Joulin, A., and Mikolov, T. (2017). Enriching Word Vectors with Subword Information. *Trans. Assoc. Comput. Linguist*. 5, 135–146. doi: 10.1162/tacl_a_00051

Carminati, M. N., Stabler, J., Roberts, A. M., and Fischer, M. H. (2006). Readers' responses to sub-genre and rhyme scheme in poetry. *Poetics* 34, 204–218. doi: 10.1016/j.poetic.2006.05.001

Castiglione, D. (2017). Difficult poetry processing: reading times and the narrativity hypothesis. *Lang. Literature* 26, 99–121. doi: 10.1177/0963947017704726

De Jaegher, H., and Di Paolo, E. (2007). Participatory sense-making: an enactive approach to social cognition. *Phenomenol. Cogn. Sci*. 6, 485–507. doi: 10.1007/s11097-007-9076-9

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *J. Exp. Psychol*. 58, 17–22. doi: 10.1037/h0046671

Dijkstra, T., and van Heuven, W. J. (2002). The architecture of the bilingual word recognition system: from identification to decision. *Bilingual. Lang. Cogn*. 5, 175–197. doi: 10.1017/S1366728902003012

Dixon, P., and Bortolussi, M. (2015). Measuring literary experience: comment on Jacobs (2016). *Sci. Study Literature* 5, 178–182. doi: 10.1075/ssol.5.2.03dix

Fitch, W. T. (2009). Prolegomena to a future science of biolinguistics. *Biolinguistics* 3, 283–320. doi: 10.4324/9780203889909

Fox, J., and Weisberg, S. (2019). *An R Companion to Applied Regression, 3rd Edn*. Thousand Oaks CA: Sage.

Gallagher, S. (2012). "Kognitionswissenschaften - Leiblichkeit und Embodiment," in *Leiblichkeit. Geschichte und Aktualität eines Konzepts*, eds E. Alloa, T. Bedorf, C. Grüny, and T. N. Klass (Tübingen: Mohr Siebeck), 320–333.

Graesser, A. C. (2015). Deeper learning with advances in discourse science and technology. *Policy Insights Behav. Brain Sci*. 2, 42–50. doi: 10.1177/2372732215600888

Grainger, J., and Jacobs, A. M. (1996). Orthographic processing in visual word recognition: a multiple read-out model. *Psychol. Rev*. 103, 518–565. doi: 10.1037/0033-295X.103.3.518

Hakemulder, F. (2020). Finding meaning through literature: 'foregrounding' as an emergent effect. *Anglistik Int. J. English Stud*. 31, 91–110. doi: 10.33675/ANGL/2020/1/8

Hakemulder, J. F. (2004). Foregrounding and its effect on readers' perception. *Discourse Process*. 38, 193–218. doi: 10.1207/s15326950dp3802_3

Hanauer, D. (1998). The genre-specific hypothesis of reading: Reading poetry and encyclopedic items. *Poetics* 26, 63–80. doi: 10.1016/S0304-422X(98)00011-4

Hartung, F., Withers, P., Hagoort, P., and Willems, R. M. (2017). When fiction is just as real as fact: No differences in reading behavior between stories believed to be based on true or fictional events. *Front. Psychol*. 8:1618. doi: 10.3389/fpsyg.2017.01618

Hintz, F., Meyer, A. S., and Huettig, F. (2020). Activating words beyond the unfolding sentence: Contributions of event simulation and word associations to discourse reading. *Neuropsychologia* 141:107409. doi: 10.1016/j.neuropsychologia.2020.107409

Hofmann, M. J., and Jacobs, A. M. (2014). Interactive activation and competition models and semantic context: from behavioral to brain data. *Neurosci. Biobehav. Rev*. 46, 85–104. doi: 10.1016/j.neubiorev.2014.06.011

Huettig, F. (2015). Four central questions about prediction in language processing. *Brain Res*. 1626, 118–135. doi: 10.1016/j.brainres.2015.02.014

Iser, W. (1972). The reading process: a phenomenological approach. *New Literary History* 3, 279–299. doi: 10.2307/468316

Iser, W. (1978). *The Act of Reading: A Theory of Aesthetic Response*. Baltimore: Johns Hopkins University Press.

Jacobs, A. M. (2011). *Neurokognitive Poetik: Elemente eines Modells des literarischen Lesens*. München: Carl Hanser Verlag.

Jacobs, A. M. (2015a). Neurocognitive poetics: methods and models for investigating the neuronal and cognitive-affective bases of literature reception. *Front. Hum. Neurosci*. 9:186. doi: 10.3389/fnhum.2015.00186

Jacobs, A. M. (2015b). The scientific study of literary experience: sampling the state of the art. *Sci. Study Literature* 5, 139–170. doi: 10.1075/ssol.5.2.01jac

Jacobs, A. M. (2015c). "Towards a neurocognitive poetics model of literary reading," in *Cognitive Neuroscience of Natural Language Use*, ed R. Willems (Cambridge: Cambridge University Press) 135–159. doi: 10.1017/CBO9781107323667.007

Jacobs, A. M. (2018). The Gutenberg English poetry corpus: exemplary quantitative narrative analyses. *Front. Digit. Human*. 5:5. doi: 10.3389/fdigh.2018.00005

Jacobs, A. M., and Kinder, A. (2015). "Worte als Worte erfahren: wie erarbeitet das Gehirn Gedichte (Experience words as words: how the brain constructs poems)," in *Kind und Gedicht (Child and Poem)*, ed A. Pompe (Berlin: Rombach), 57–76.

Jacobs, A. M., and Kinder, A. (2017). "The brain is the prisoner of thought": a machine-learning assisted quantitative narrative analysis of literary metaphors for use in neurocognitive poetics. *Metaphor Symbol* 32, 139–160. doi: 10.1080/10926488.2017.1338015

Jacobs, A. M., Lüdtke, J., Aryani, A., Meyer-Sickendieck, B., and Conrad, M. (2016). Mood-empathic and aesthetic responses in poetry reception. *Sci. Study Literature* 6, 87–130. doi: 10.1075/ssol.6.1.06jac

Jacobs, A. M., V o, M. L., Briesemeister, B. B., Conrad, M., Hofmann, M. J., Kuchinke, L., et al. (2015). 10 years of BAWLing into affective and aesthetic processes in reading: what are the echoes? *Front. Psychol*. 6:714. doi: 10.3389/fpsyg.2015.00714

Jacobs, A. M., and Willems, R. M. (2018). The fictive brain: neurocognitive correlates of engagement in literature. *Rev. Gen. Psychol*. 22, 147–160. doi: 10.1037/gpr0000106

Jakobson, R. (1960). "Concluding Statement: Linguistics and Poetics," in *Style in Language*, ed T. Sebeok (Cambridge: MIT Press), 350–377.

Jolsvai, H., Mccauley, S. M., and Christiansen, M. H. (2013). "Meaning overrides frequency in idiomatic and compositional multiword chunks," in *Proceedings of the 35th Annual Meeting of the Cognitive Science Society*, 692–697. Available online at: https://escholarship.org/uc/item/5cv7b5xs#author

Kao, J., and Jurafsky, D. (2012). "A computational analysis of style, affect, and imagery in contemporary poetry," in *Proceedings of the NAACL-HLT 2012 Workshop on Computational Linguistics for Literature* (Montréal, QC: Association for Computational Linguistics), 8–17.

Kentner, G. (2012). Linguistic rhythm guides parsing decisions in written sentence comprehension. *Cognition* 123, 1–20. doi: 10.1016/j.cognition.2011.11.012

Kiss, G. R., Armstrong, C., Milroy, R., and Piper, J. (1973). "An associative thesaurus of English and its computer analysis," in *The Computer and Literary Studies*, eds A. J. Aitkin, R. W. Bailey, and N. Hamilton-Smith (Edinburgh: University Press), 153–165.

Kuchinke, L., Jacobs, A. M., Grubich, C., Vø, M. L., Conrad, M., and Herrmann, M. (2005). Incidental effects of emotional valence in single word processing: an fMRI study. *Neuroimage* 28, 1022–1032. doi: 10.1016/j.neuroimage.2005.06.050

Kuiken, D., Campbell, P., and Sopčák, P. (2012). The experiencing questionnaire. *Sci. Study Literature* 2, 243–272. doi: 10.1075/ssol.2.2.04kui

Lüdtke, J., Meyer-Sickendieck, B., and Jacobs, A. M. (2014). Immersing in the stillness of an early morning: Testing the mood empathy hypothesis of poetry reception TT - Eintauchen in die Stille eines frühen Morgens: Testen der

Stimmungsempathie-Hypothese der Poesiewahrnehmung. *Psychol. Aesthet. Creativity Arts* 8, 363–377. doi: 10.1037/a0036826

Magezi, D. A. (2015). Linear mixed-effects models for within-participant psychology experiments: an introductory tutorial and free, graphical user interface (LMMgui). *Front. Psychol.* 6:2. doi: 10.3389/fpsyg.2015.00002

Magyari, L., Mangen, A., Kuzmičová, A., Jacobs, A. M., and Lüdtke, J. (2020). Eye movements and mental imagery during reading of literary texts in different narrative styles. *J. Eye Movement Res.* 13:3. doi: 10.16910/jemr.13.3.3

Mangan, B. (2001). Sensation's ghost: the non-sensory "fringe" of consciousness. *Psyche* 7.

Mangan, B. (2008). Representation, rightness and the fringe. *J. Conscious. Stud.* 15, 75–82. Available online at: http://www.cogprints.org/7591

Menninghaus, W., Wagner, V., Wassiliwizky, E., Jacobsen, T., and Knoop, C. A. (2017). The emotional and aesthetic powers of parallelistic diction. *Poetics* 63, 47–59. doi: 10.1016/j.poetic.2016.12.001

Mukarovský, J. (1964). "Standard language and poetic language," in *A Prague School Reader on Esthetics, Literary Structure, and Style*, ed P. L. Garvin (Washington, DC: Georgetown University Press), 17–30.

Müller, H. J., Geyer, T., Günther, F., Kacian, J., and Pierides, S. (2017). Reading English-language haiku: processes of meaning construction revealed by eye movements. *J. Eye Movement Res.* 10, 1–33. doi: 10.16910/10.1.4

Murray, W. S., and Forster, K. I. (2004). Serial mechanisms in lexical access: the rank hypothesis. *Psychol. Rev.* 111, 721–756. doi: 10.1037/0033-295X.111.3.721

Nakagawa, S. and Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods Ecol. Evol.* 4, 133–142. doi: 10.1111/j.2041-210x.2012.00261.x

Netzer, Y., Gabay, D., Goldberg, Y., and Elhadad, M. (2009). "Gaiku: generating Haiku with word associations norms," in *Proceedings of the NAACL HLT Workshop on Computational Approaches to Linguistics Creativity*, 32–39. doi: 10.3115/1642011.1642016

Obermeier, C., Menninghaus, W., von Koppenfels, M., Raettig, T., Schmidt-Kassow, M., Otterbein, S., et al. (2013). Aesthetic and emotional effects of meter and rhyme in poetry. *Front. Psychol.* 4:10. doi: 10.3389/fpsyg.2013.00010

Osowiecka, M., and Kolanczyk, A. (2018). Let's read a poem! What type of poetry boosts creativity? *Front. Psychol.* 9:1781. doi: 10.3389/fpsyg.2018.01781

Panksepp, J. (2008). The power of the word may reside in the power of affect. *Integr. Psychol. Behav. Sci.* 42, 47–55. doi: 10.1007/s12124-007-9036-5

Papp-Zipernovszky, O., Mangen, A., Jacobs, A., and Lüdtke, J. (2021). Shakespeare sonnets reading: an empirical study of emotional responses. *Lang. Literature*.

Paulmann, S., Bleichner, M., and Kotz, S. A. (2013). Valence, arousal, and task effects in emotional prosody processing. *Front. Psychol.* 4:345. doi: 10.3389/fpsyg.2013.00345

Peelle, J. E., and Davis, M. H. (2012). Neural oscillations carry speech rhythm through to comprehension. *Front. Psychol.* 3:320. doi: 10.3389/fpsyg.2012.00320

Pennington, J., Socher, R., and Manning, C. D. (2014). "GloVe: global vectors for word representation," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (Doha: Association for Computational Linguistics), 1532–1543. doi: 10.3115/v1/D.14-1162

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Reber, R., Fazendeiro, T. A., and Winkielman, P. (2002). Processing fluency as the source of experiences at the fringe of consciousness. *Psyche* 8.

Reber, R., Schwarz, N., and Winkielman, P. (2004). Processing fluency and aesthetic pleasure: is beauty in the perceiver's processing experience? *Pers. Soc. Psychol. Rev.* 8, 364–382. doi: 10.1207/s15327957pspr0804_3

Řehůřek, R., and Sojka, P. (2010). "Software framework for topic modelling with large corpora," in *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks* (Valletta: ELRA), 45–50.

Roediger, H. L., and McDermott, K. B. (1995). Creating false memories: remembering words not presented in lists. *J. Exp. Psychol. Learn. Memory Cogn.* 21, 803–814. doi: 10.1037/0278-7393.21.4.803

Rosenblatt, L. M. (1978). *The Reader, the Text, the Poem: The Transactional Theory of the Literary Work*. Carbondale, IL: Southern Illinois University Press.

Sluckin, W., Colman, A. M., and Hargreaves, D. J. (1980). Liking words as a function of the experienced frequency of their occurrence. *Brit. J. Psychol.* 71, 163–169. doi: 10.1111/j.2044-8295.1980.tb02742.x

Stadthagen-Gonzalez, H., and Davis, C. J. (2006). The Bristol norms for age of acquisition, imageability, and familiarity. *Behav. Res. Methods* 38, 598–605. doi: 10.3758/BF03193891

Sylvester, T., Braun, M., Schmidtke, D., and Jacobs, A. M. (2016). The Berlin Affective Word List for children (kidBAWL): exploring processing of affective lexical semantics in the visual and auditory modalities. *Front. Psychol.* 7:969. doi: 10.3389/fpsyg.2016.00969

Teng, D. W., Wallot, S., and Kelty-Stephen, D. G. (2016). Single-word recognition need not depend on single-word features: narrative coherence counteracts effects of single-word features that lexical decision emphasizes. *J. Psycholinguist. Res.* 45, 1451–1472. doi: 10.1007/s10936-016-9416-4

Topping, K. J. (2015). Fiction and non-fiction reading and comprehension in preferred books. *Read. Psychol.* 36, 350–387. doi: 10.1080/02702711.2013.865692

Van Casteren, M., and Davis, M. H. (2006). Mix, a program for pseudorandomization. *Behav. Res. Methods* 38, 584–589. doi: 10.3758/BF03193889

van Heuven, W. J., Mandera, P., Keuleers, E., and Brysbaert, M. (2014). SUBTLEX-UK: a new and improved word frequency database for British English. *Q. J. Exp. Psychol.* 67, 1176–1190. doi: 10.1080/17470218.2013.850521

Van Krieken, K., Hoeken, H., and Sanders, J. (2015). From reader to mediated witness: the engaging effects of journalistic crime narratives. *J. Mass Commun. Q.* 92, 580–596. doi: 10.1177/1077699015586546

Venables, W. N., and Ripley, B. D. (2002). *Modern Applied Statistics with S, 4th Edn*. New York, NY: Springer. doi: 10.1007/978-0-387-21706-2

Warriner, A. B., Kuperman, V., and Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behav. Res. Methods* 45, 1191–1207. doi: 10.3758/s13428-012-0314-x

Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt. II. *Psychol. Forsch.* 4, 301–350. doi: 10.1007/BF00410640

Westphal-Fitch, G., and Fitch, W. T. (2018). *Bioaesthetics: The Evolution of Aesthetic Cognition in Humans and Other Animals*, Vol. 237. Elsevier. doi: 10.1016/bs.pbr.2018.03.003

Winkielman, P., Niedenthal, P., Wielgosz, J., Eelen, J., and Kavanagh, L. C. (2014). "Embodiment of cognition and emotion," in *APA Handbook of Personality and Social Psychology, Volume 1: Attitudes and Social Cognition*, eds M. Mikulincer, P. R. Shaver, E. Borgida, and J. A. Bargh (Washington, DC: American Psychological Association), 151–175. doi: 10.1037/14341-004

Wolfe, M. B. (2005). Memory for narrative and expository text: Independent influences of semantic associations and text organization. *J. Exp. Psychol. Learn. Memory Cogn.* 31, 359–364. doi: 10.1037/0278-7393.31.2.359

Xue, S., Jacobs, A. M., and Lüdtke, J. (2020). What is the difference? Rereading Shakespeare's sonnets –an eye tracking study. *Front. Psychol.* 11:421. doi: 10.3389/fpsyg.2020.00421

Zajonc, R. B. (2000). *Feeling and Thinking: Closing the Debate Over the Independence of Affect*. New York, NY: Cambridge University Press.

Zwaan, R. A. (1994). Effect of genre expectations on text comprehension. *J. Exp. Psychol. Learn. Memory Cogn.* 20, 920–933. doi: 10.1037/0278-7393.20.4.920

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.

frontiers in Psychology

BRIEF RESEARCH REPORT
published: 26 October 2021
doi: 10.3389/fpsyg.2021.721783

Can Masked Emotion-Laden Words Prime Emotion-Label Words? An ERP Test on the Mediated Account

Chenggang Wu[1,2], *Juan Zhang*[2,3*] *and Zhen Yuan*[3,4]

[1] *Multilingual Edu-AI Laboratory, School of Education, Shanghai International Studies University, Shanghai, China,* [2] *Faculty of Education, University of Macau, Macau, China,* [3] *Centre for Cognitive and Brain Sciences, University of Macau, Macau, China,* [4] *Faculty of Health Sciences, University of Macau, Macau, China*

The present event-related potential (ERP) study explored whether masked emotion-laden words could facilitate the processing of both emotion-label words and emotion-laden words in a valence judgment task. The results revealed that emotion-laden words as primes failed to influence target emotion-label word processing, whereas emotion-laden words facilitated target emotion-laden words in the congruent condition. Specifically, decreased late positivity complex (LPC) was elicited by emotion-laden words primed by emotion-laden words of the same valence than those primed by emotion-laden words of different valence. Nevertheless, no difference was observed for emotion-label words as targets. These findings supported the mediated account that claimed emotion-laden words engendered emotion *via* the mediation of emotion-label words and hypothesized that emotion-laden words could not prime emotion-label words in the masked priming paradigm. Moreover, this study provided additional evidence showing the distinction between emotion-laden words and emotion-label words.

Keywords: emotion-label words, emotion-laden words, emotion words, ERP, LPC

OPEN ACCESS

Edited by:
Francesca M. M. Citron,
Lancaster University, United Kingdom

Reviewed by:
Yanyun Zhou,
Nanjing University of Posts and Telecommunications, China
José Antonio Hinojosa,
Complutense University of Madrid, Spain

****Correspondence:***
Juan Zhang
juanzhang@um.edu.mo

Specialty section:
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

Received: *07 June 2021*
Accepted: *21 September 2021*
Published: *26 October 2021*

Citation:
Wu C, Zhang J and Yuan Z (2021) Can Masked Emotion-Laden Words Prime Emotion-Label Words? An ERP Test on the Mediated Account. Front. Psychol. 12:721783. doi: 10.3389/fpsyg.2021.721783

INTRODUCTION

Across various studies of different languages, it has been consistently found that emotion-label words and emotion-laden words differ in a variety of tasks, such as lexical decision task (Kazanas and Altarriba, 2015, 2016a,b; Zhang et al., 2017, 2018a), flanker task (Wu and Zhang, 2019a; Zhang et al., 2019a,b), and affective Simon task (Altarriba and Basnight-Brown, 2011). Emotion-label words (e.g., shame and ecstasy) directly describe an emotional state, whereas emotion-laden words (e.g., butterfly and surgery) indirectly induce emotion *via* elaboration (Sutton and Altarriba, 2016; Zhou and Tse, 2020). Recent event-related potential (ERP) studies found that more substantial brain activation was evoked by emotion-label words than emotion-laden words in both Chinese (Zhang et al., 2017; Wang et al., 2019) and English (Zhang et al., 2018a) in a lexical decision task. For example, Zhang et al. (2018a) found that emotion-label words in English provoked larger N170 than emotion-laden words. There was also research demonstrating that the discrepancy between emotion-laden words and emotion-label words had an impact on perception of emotion (Wu et al., 2019, 2020). Affective picture valence judgment was facilitated by preceding emotion-label words over emotion-laden words with accentuated processing speed and weaker electrophysiological responses, and the facilitation effect was found in both Chinese (Wu et al., 2020) and English (Wu et al., 2019).

Although the separation between emotion-laden words and emotion-label words has received much support (Wu and Zhang, 2019b), it is still unclear how the two kinds of emotion words relate to each other. Emotion-label words and emotion-laden words are by no means irrelevant because when recognizing emotion-laden words, individuals will naturally activate related emotion-label words accordingly (Wu and Zhang, 2019b). For example, when individuals read the word "reward," they will be reminded of the experiences of receiving a reward and emotions that are embodied in those experiences will also be induced. However, the connection between emotion-label words and emotion-laden words is not one-to-one. Negative emotion-laden words, such as death, will activate various negative emotion concepts, such as fear, sadness, and so on. Therefore, one emotion-laden word can have multiple connections to many emotion-label words, but each of the connections is dependent on situations. The current study employed the masked priming paradigm to examine whether or not masked emotion-laden words could prime emotion-label words.

Investigating the association between emotion-laden words and emotion-label words is of close relevance to the theory explaining the distinction between emotion-label words and emotion-laden words (Knickerbocker, 2014). Altarriba and Basnight-Brown (2011) proposed a mediated account to explain the differences among emotion-label words and emotion-laden words and argued that emotion-laden words could be regarded as a type of "mediated" emotion concepts. Unlike emotion-label words that label emotion concepts straightforwardly, emotion-laden words elicit emotion mediated by the emotion-label words after emotional experiences related to the emotion-laden words are elaborated. Therefore, emotion-label words generated increased emotion activation than emotion-laden words, and this finding has been widely reported (Knickerbocker and Altarriba, 2013; Kazanas and Altarriba, 2015, 2016a,b; Zhang et al., 2017, 2018a, 2019a,b; Wang et al., 2019; Wu et al., 2019, 2020). However, the mediated account claimed by Altarriba and Basnight-Brown (2011) does not specify how emotion-laden words are mediated by emotion-label words. As elucidated before, one emotion-laden word does not activate a single and certain emotion labeled by one emotion-label word. Rather, each emotion-laden word has oblique relationships with emotion activation *via* unpredictable connections to emotion-label words. This individualized and contextualized mapping of emotion-laden words and emotion-label words is an important addition to the original mediated account (Altarriba and Basnight-Brown, 2011).

Extant studies have already provided partial support for the mediated account. Kazanas and Altarriba (2015) examined how emotion-label words and emotion-laden words were different in provoking priming effects in masked and unmasked priming paradigms. In their study, emotion-label words primed emotion-label words, and emotion-laden words primed emotion-laden words. The results revealed a significant priming effect for both emotion-label words and emotion-laden words with the priming effect being larger for emotion-label words, suggesting that emotion-label words and emotion-laden words are distinct categories.

Although Kazanas and Altarriba (2015, 2016a,b) offered much insight into how emotion-label words and emotion-laden words are semantically presented, they did not provide a certain answer to how emotion-laden words and emotion-label words are related to each other. In other words, in their study, when target words were emotion-label words, the primes were also emotion-label words, and the same procedure was applied to emotion-laden words. This operation resulted in one unresolved issue whether emotion-laden words could prime emotion-label words. According to the mediated account (Altarriba and Basnight-Brown, 2011), if emotion-laden words elicit emotion *via* emotion concepts that are labeled by emotion-label words in an individualized and a contextualized manner, it can thus be predicted that emotion-laden words in the masked priming paradigm would not prime emotion-label words as targets. This study attempted to examine whether masked emotion-laden words could prime emotion-label words and test the hypothesis derived from the mediated account (Knickerbocker et al., 2019). As suggested by Knickerbocker et al. (2019), one extension of the study by Kazanas and Altarriba (2015) was to investigate how emotion-label words are primed by emotion-laden words.

One recent event-related potential (ERP) study (Wu et al., 2021) explicitly examined how emotion-label words and emotion-laden words as primes influence target emotion-laden word processing in both masked and unmasked priming paradigms. The overall results confirmed the division of the two types of words, and, more importantly, this distinction could still be observed in the masked condition. Specifically, masked emotion-label words inhibited target emotion-laden words by increasing the error rate and decreasing the processing speed than those target words preceded by masked emotion-laden words. However, one unresolved problem was how emotion-laden words could influence emotion-label words. This study aimed to answer this question according to a previous study (Wu et al., 2021).

In addition, this study measured electrophysiological responses using the ERP technique. One late ERP component (late positivity complex, LPC) related to elaboration during emotion word processing was explored in this study. Emotion-laden words were found to elicit enhanced LPC than emotion-label words, suggesting that processing emotion-laden words is more effortful than emotion-label words (Zhang et al., 2018a). If emotion-label words could be primed by emotion-laden words in the masked condition, it is predicted that emotion-label words would generate a larger LPC when the emotion-label words are preceded by unrelated emotion-laden words (from the different valence) than by related emotion-laden words (from the same valence). If emotion-laden words could not facilitate emotion-label word processing, no modulation on LPC is expected.

METHOD

Participants

About 25 Chinese-English bilingual speakers from the University of Macau were recruited for this study. Due to exceeding artifacts, 5 were excluded, and the remaining 20 Chinese speakers were

TABLE 1 | Mean and SD in brackets for word characteristics for Chinese emotion-laden words as primes.

	Emotion-laden words Primed by label words		Emotion-laden words Primed by laden words	
	Negative	**Positive**	**Negative**	**Positive**
Sample	死囚 (prisoner)	花束 (flower)	火化 (cremation)	春光 (Spring view)
Word frequency	1.98 (0.79)	1.97 (0.69)	2.14 (0.64)	1.82 (0.75)
Strokes	17.43 (3.94)	18.08 (4.31)	17.70 (4.82)	16.78 (4.86)
Arousal	5.58 (1.36)	5.52 (1.03)	5.43 (1.03)	5.40 (0.91)
Valence	3.49 (1.07)	5.91 (0.63)	3.17 (0.84)	5.98 (0.60)

TABLE 2 | Mean and SD in brackets for word characteristics for Chinese emotion-label words and emotion-laden words as targets.

	Emotion-label words		Emotion-laden words	
	Negative	**Positive**	**Negative**	**Positive**
Sample	心疼 (sadness)	舒心 (joy)	分手 (breakup)	生日 (birthday)
Word frequency	2.37 (0.86)	2.22 (0.84)	2.30 (0.67)	2.37 (0.63)
Strokes	16.78 (4.24)	17.80 (3.31)	16.23 (4.31)	16.55 (5.04)
Arousal	4.78 (0.29)	4.78 (0.47)	4.75 (0.45)	4.88 (0.67)
Valence	2.58 (0.32)	5.29 (0.48)	2.32 (0.57)	5.34 (0.51)

kept for further processing (3 men, mean age: 27 years). All of the participants were right-handed. They reported that they did not suffer from psychiatric disorders or brain damage. Participants had normal or corrected-to-normal vision. The sample size was determined by calculating the prior power using G*power (Faul et al., 2007). Repeated-measure ANOVA requires at least 20 participants when the power is 0.8, and effect size is medium, partial $\eta^2 = 0.1$, also in line with previous studies (Wu and Zhang, 2019a; Zhang et al., 2019a).

Stimuli

Two sets of Chinese emotion words were selected for the stimuli (Wu et al., 2021). The first set of emotion-laden words that were used as primes was obtained from a recent Chinese norming database (Yao et al., 2017). There were 160 emotion-laden words, including 80 positive words and 80 negative words. Both the negative and positive words were divided into two halves to prime 160 target emotion words. The target words formed the second set of Chinese words, including 80 emotion-label words (40 positive and 40 negative words) and 80 emotion-laden words (40 positive and 40 negative words). The primes were not different among the different conditions in terms of word frequency, strokes, and arousal, all $ps > 0.05$. The same restriction was applied to target words, all $ps > 0.05$ (see **Tables 1**, **2** for more details). Primes and targets were randomly combined to control the semantic association between the primes and targets. We calculated the word association strength between emotion-label words and emotion-laden words from a recent Chinese word database (Lin et al., 2019) and found no association between the primes and targets (see more details in Discussion).

Procedure

All procedures have been approved by the Institutional Review Board at the University of Macau before the experiment. All participants first signed a consent form before they started the experiment. After setting up the Geodesic Sensor Net (EGI, Eugene, OR, USA), the experimenter described the experimental procedure to participants. At the same time, participants could read the written instruction that was displayed on the monitor at a distance of 70 cm. The task was to determine the valence of the target Chinese emotion words. Each trial started with a 500 ms fixation. A forward mask lasting for 500 ms was followed by a prime word (50 ms). The prime was also masked by a backward mask (10 ms) that was created by overlapping several complex Chinese characters (Zhang et al., 2018b). Afterward, a target emotion word (Song font, 48 points) was presented to participants and disappeared as soon as participants made a response. Each prime was presented two times in two different blocks and was randomly paired with a target word with the same (related condition) or different (unrelated condition) valence. Altogether, 320 trials were dispersed into 8 blocks, each of which included 40 trials. Trials within each block and the blocks were presented randomly. The short rests were inserted between the blocks.

ERP Recording and Analysis

Scalp voltages were recorded with a 129-channel Geodesic Sensor Net with a sampling rate of 1,000 Hz. The impedance was retained below 50 kΩ during recording. To avoid eyeblinks, at the end of each trial, a notice of allowing eyeblink was displayed for 1,000 ms. The offline data were first filtered with a bandpass of 0.1–30 Hz. The continuous EEG data were further segmented into epochs from 100 ms prior to the presence of the target words. The segmentations were passed through an artifact scan (eyeblink, 70 μV; eye movement, 27.57 μV) and were discarded if the epochs were labeled as an eyeblink or eye movement. The channel was labeled as bad if the change of amplitude exceeded 200 μV. The bad channels were replaced by peripherical sites. However, we deleted the segmentations with more than 10% bad channels. The EEG data were referenced to the average of all electrodes. A baseline correction of −100 to 0 ms on the onset of stimuli was performed. For LPC, three electrodes (Cz, C2, and C4) were chosen during the time window of 500–800 ms. LPC was mostly identified around central sites (Zhang et al., 2018a). We also determine the electrodes using a visual inspection. Moreover, previous studies also found LPC was more salient in the right hemisphere than in the left hemisphere (Zhang et al., 2018a), thus the right central sites were chosen for LPC in this study.

RESULTS

Behavioral Results

We deleted the trials that participants made responses slower or faster than Mean ± 2.5 SD for the behavioral data

TABLE 3 | Mean reaction time (ms) and accuracy rate (%) of emotion-label words and emotion-laden words in brackets as a function of relatedness and valence.

	Negative		Positive	
	Related	**Unrelated**	**Related**	**Unrelated**
Emotion-label	734.17 (95.25)	728.50 (95.63)	720.47 (96.88)	714.17 (96.25)
Emotion-laden	711.53 (95.75)	718.75 (96.75)	669.92 (99.25)	688.37 (98.75)

analysis, thereby missing 2.19% data. A three-way ANOVA was conducted. The three within-subject factors were emotion word type, valence, and relatedness.

The behavioral results are summarized in **Table 3**. Target emotion-label words (0.96) were recognized less accurately than emotion-laden words (0.98), [$F_{(1, 19)} = 19.343$, $p < 0.001$, partial $\eta^2 = 0.504$]. The accuracy rate was higher for positive words (0.98) than for negative words (0.96), [$F_{(1, 19)} = 9.615$, $p < 0.01$, partial $\eta^2 = 0.336$]. There was no significant main effect of relatedness and other interactions in the accuracy rate analysis, $ps > 0.05$. Regarding reaction times, emotion-laden words (697.14 ms) were processed faster than emotion-label words (724.33 ms), [$F_{(1, 19)} = 10.016$, $p < 0.01$, partial $\eta^2 = 0.345$]. In addition, negative words (723.24 ms) were recognized slower than positive words (698.23 ms), [$F_{(1, 19)} = 7.623$, $p < 0.05$, partial $\eta^2 = 0.286$]. No other main effects or interactions were found to reach significance, $ps > 0.5$.

ERP Results

We performed a 2 (Valence: negative and positive) × 2 (Relatedness: related and unrelated) × 2 (Emotion word type: emotion-laden words and emotion-label words) repeated-measure ANOVA. The priming effect was confirmed such that emotion words in the related condition (0.58 μV) provoked a smaller LPC than in the unrelated condition (0.97 μV), [$F_{(1, 19)} = 5.732$, $p < 0.05$, partial $\eta^2 = 0.232$]. To further compare the priming effect between emotion-label words and emotion-laden words, additional ANOVAs containing the two within-subject factors (valence and relatedness) for emotion-label words and emotion-laden words were conducted separately. The result showed that emotion-laden words only produced a priming effect on emotion-laden words, [$F_{(1, 19)} = 4.585$, $p < 0.05$, partial $\eta^2 = 0.194$], rather than on emotion-label words, [$F_{(1, 19)} = 1.327$, $p > 0.1$]. A larger LPC was elicited by the target emotion-laden words that were preceded by the different valence emotion-laden words (1.17 μV) than those that were preceded by the same valence emotion-laden words (0.64 μV). However, no priming effect was found for emotion-label words as targets (0.52 μV in the related condition and 0.77 μV in the unrelated condition). No other main effects or interactions were identified, $ps > 0.05$ (refer to **Figure 1**).

DISCUSSION

In this experiment, we investigated whether emotion-laden words as primes could influence target emotion-label and emotion-laden words in the masked priming paradigm. The behavioral results showed the processing differences between the two types of emotion words, replicating many prior examinations (Kazanas and Altarriba, 2015, 2016a,b; Zhang et al., 2017, 2018a, 2019a,b; Wang et al., 2019; Wu and Zhang, 2019a,b; Wu et al., 2019, 2020). Electrophysiological evidence further supported the emotion word-type effect by showing that emotion-laden words rather than emotion-label words could be facilitated by emotion-laden words as masked primes, suggesting that the two categories are emotion-label words and emotion-laden words (Wu and Zhang, 2019b).

This study aimed to examine the mediated account that explained the relationship and the differences between emotion-label words and emotion-laden words (Altarriba and Basnight-Brown, 2011). Altarriba and Basnight-Brown (2011) argued that emotion-laden words are "mediated" by the emotion concepts that are indicated by emotion-label words. Therefore, the emotion activation that is induced by emotion-laden words is achieved after related emotion concepts are activated. For example, a recent study showed that emotion-laden words generated a more substantial electrophysiological activation than emotion-label words, indicating that emotion-laden word processing is more effortful than emotion-label words (Zhang et al., 2018a). However, at a first glance, this study showed contradictory evidence that emotion-laden words had a higher processing speed than emotion-label words with an increased accuracy rate, implying that emotion-label words were harder to recognize in a valence judgment task. This result pattern was, indeed, in line with one recent study that used the emotion flanker task and found that Chinese emotion-laden words were recognized faster than emotion-label words (Zhang et al., 2019b). One difference between the two studies was that there were only 6 emotion-label words in each category in that study (Zhang et al., 2019b), but there were 80 emotion-label words in this study. Therefore, this study, using a large number of emotion words, was an extension of a previous study (Zhang et al., 2019b) by showing that judging the valence of emotion-label words was hard in both flanker and affective priming tasks.

The reason for the difficulty in evaluating the valence of the emotion-label words is that most emotion concepts that are labeled by emotion-label words are not a valence-based representation, especially for negative emotions (e.g., shame, sadness, fear, anger, and boredom). Many researchers (Ekman, 1992; Izard, 2007; Ekman and Cordaro, 2011; Lench et al., 2011) theorized that negative emotions were discrete and negativity was not sufficient to explain the variance between the negative emotions, such as fear and sadness. The ambiguous valence conveyed by emotion-label words was to fill an important void in the mediated account explaining the association between emotion-label words and emotion-laden words. The mediated account claimed that emotion-laden words were mediated by emotion-label words but did not specify how emotion-laden words were mediated by emotion-label words. The emotion-laden words failed to prime emotion-label words, suggesting that emotion-laden words were mediated by emotion-label words in ambiguous associations between the two kinds of emotion words. For example, the word "wedding" can induce many related positive emotions (e.g., happiness and excitement). More importantly, emotions are discrete in both negative and positive

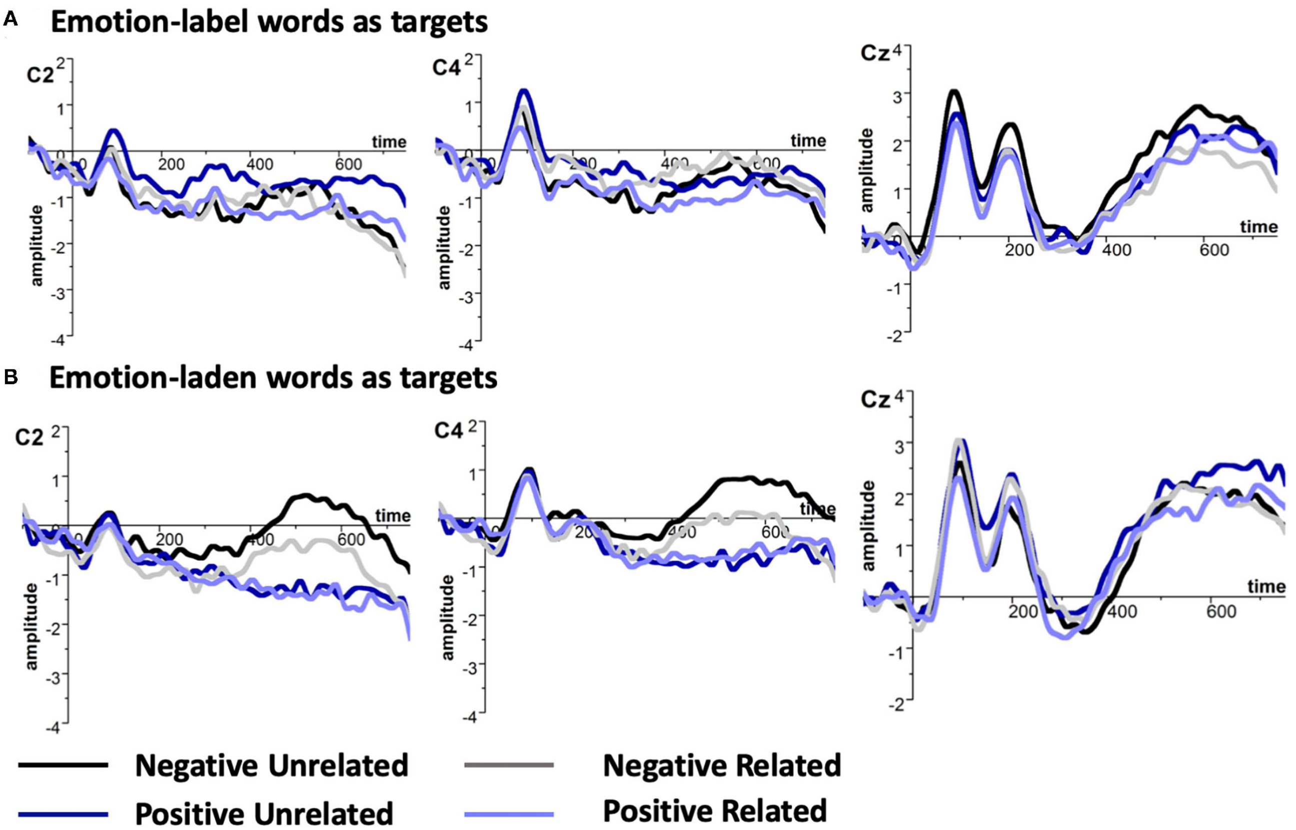

FIGURE 1 | Grand average event-related potentials (ERPs) of late positivity complex (LPC) at selected electrodes **(A)** emotion-label words and **(B)** emotion-laden words as primes, amplitude (μV), and time (ms).

categories (Lench et al., 2011; Shiota et al., 2017), increasing the difficulty of judging the valence of target emotion-label words.

It could be argued that word association might influence the priming effect (Hines et al., 1986). We controlled the word association between primes and targets by randomly pairing primes and targets. Therefore, it is assumed that the word association between primes and targets is nearly zero. Based on the Chinese Lexical Association Database (CLAD), one recent Chinese association norming database (Lin et al., 2019), we further analyzed the word association between primes and targets. We retrieved the Baroni-Urbani measure on clauses for each prime emotion word and found that almost all the primes were not associated to targets both in related and unrelated conditions, except a very few words [see the **Appendix (Supplementary Material)** for the word list and word association, and the two prime words were not found in CLAD]. Further comparisons between primes and targets in related and unrelated for emotion-label words and emotion-laden words found that the word association strength was equally weak for emotion-label words and emotion-laden words as targets, [$F_{(1, 77)} = 0.737$, $p > 0.39$], for emotion word type, [$F_{(1, 77)} = 1.572$, $p > 0.21$], for relatedness, and [$F_{(1, 77)} = 0.145$, $p > 0.70$] for the interaction between emotion word type and relatedness. Based on the restricted word association between primes and targets, it is clear that the priming effect for emotion-laden words was not semantic but affective in essence.

Several limitations of this study should be noted. The first limitation is the definition of emotion-label words and emotion-laden words. It is argued that research on emotion-label words and emotion-laden words is lacking an objective measurement of determining what is an emotion-label word or an emotion-laden word (Hinojosa et al., 2020). One recent normative database of 1,286 Spanish words proffered the ratings of emotional prototypicality that refers to the degree of the typicality of an emotion word (Pérez et al., 2021). The higher prototypicality of an emotion word means that it is more reasonable to be defined as an emotion-label word, such as fear. However, this study did not use this approach to define emotion-label words. Future studies can use the prototype approach to objectively define an emotion-label word. The second limitation is that only adults were included in this study. There is a recent urge to explore how emotional concepts are acquired by children (Hoemann et al., 2019). Therefore, how children process emotion-label words and emotion-laden words would enlighten the emotional development of children across cultures. Future exploration could follow this trend to differentiate emotion-label words and emotion-laden words and examine how emotion-laden words and emotion-label words are associated in mental lexicon of children. The third limitation is that we did not control the concreteness between emotion-label words and emotion-laden words. We retrieved the concreteness of primes and targets from a recent Chinese normative database (Xu and Li, 2020) and found that emotion-label words (48 words were found) were more abstract than emotion-laden words (60 words were found), [$F_{(1,106)} = 61.426, p < 0.001$]. Therefore, the distinction between the two kinds of words can be attributed to the influence of concreteness. However, the primes (emotion-laden words) for target emotion-label words (55 prime words were found) and emotion-laden words (45 prime words were found) were the same on concreteness, [$F_{(1, 98)} < 1, p > 0.47$]. The priming effect was found only for emotion-laden word pairs, and this result thus is not related to concreteness. Further research on emotion-label words and emotion-laden words should consider controlling concreteness (Wang et al., 2019). In addition, although we used the masked affective priming paradigm to preclude the strategic influence of processing prime words, future studies can also exploit the unmasked priming paradigm and vary the duration of primes (extending to 1,000 ms) to explore the affective priming of the two types of words (Kazanas and Altarriba, 2016a).

To summarize, the results in this study supported the mediated account, assuming that there was no priming effect of emotion-laden words on emotion-label words. However, the priming effect of emotion-laden words on emotion-laden target words was identified, in line with the mediated account and previous studies (Kazanas and Altarriba, 2015, 2016a,b). Moreover, the distinction between the two kinds of emotion words was also replicated, compatible with the emotion conflict study using the flanker task (Zhang et al., 2019b).

DATA AVAILABILITY STATEMENT

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Institutional Review Board in the University of Macau. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

CW and JZ developed the research idea and determined the research design. ZY commented on the research design critically. CW conducted the research, analyzed the data, and drafted the manuscript. JZ and ZY reviewed the manuscript and provided insightful revisions at all stages. All authors have read and agreed to the published version of the manuscript.

FUNDING

This study was funded by the University of Macau (MYRG2017-00217-FED, MYRG2016-00193-FED, and MYRG2015-00221-FED). This study was also supported by the Shanghai Philosophy and Social Science Planning Youth Project (2020EYY004), Shanghai Chenguang Talent Program (20CG40), and the Innovative Research Team of Shanghai International Studies University (2020114052).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2021.721783/full#supplementary-material

REFERENCES

Altarriba, J., and Basnight-Brown, D. M. (2011). The representation of emotion vs. emotion-laden words in English and Spanish in the affective simon task. *Int. J. Bilingual.* 15, 310–328. doi: 10.1177/1367006910379261

Ekman, P. (1992). An argument for basic emotions. *Cogn. Emot.* 6, 169–200. doi: 10.1080/02699939208411068

Ekman, P., and Cordaro, D. (2011). What is meant by calling emotions basic. *Emotion Rev.* 3, 364–370. doi: 10.1177/1754073911410740

Faul, F., Erdfelder, E., Lang, A.-G., and Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavior, and biomedical sciences. *Behav. Res. Methods Instrum. Comput.* 39, 175–191. doi: 10.3758/BF03193146

Hines, D., Czerwinski, M., Sawyer, P. K., and Dwyer, M. (1986). Automatic semantic priming: effect of category exemplar level and word association level. *J. Exp. Psychol. Hum. Percept. Perform.* 12, 370–379. doi: 10.1037/0096-1523.12.3.370

Hinojosa, J. A., Moreno, E., and Ferré, P. (2020). On the limits of affective neurolinguistics: a "universe" that quickly expands. *Lang. Cogn. Neurosci.* 35, 1–8. doi: 10.1080/23273798.2020.1761988

Hoemann, K., Xu, F., and Barrett, L. F. (2019). Emotion words, emotion concepts, and emotional development in children: a constructionist hypothesis. *Dev. Psychol.* 55, 1830–1849. doi: 10.1037/dev0000686

Izard, C. E. (2007). Basic emotions, natural kinds, emotion schemas, and a new paradigm. *Perspect. Psychol. Sci.* 2, 260–280. doi: 10.1111/j.1745-6916.2007.00044.x

Kazanas, S. A., and Altarriba, J. (2015). The automatic activation of emotion and emotion-laden words: evidence from a masked and unmasked priming paradigm. *Am. J. Psychol.* 128, 323–336. doi: 10.5406/amerjpsyc.128.3.0323

Kazanas, S. A., and Altarriba, J. (2016a). Emotion word processing: effects of word type and valence in Spanish–English bilinguals. *J. Psycholinguist. Res.* 45, 395–406. doi: 10.1007/s10936-015-9357-3

Kazanas, S. A., and Altarriba, J. (2016b). Emotion word type and affective valence priming at a long stimulus onset asynchrony. *Lang. Speech* 59, 339–352. doi: 10.1177/0023830915590677

Knickerbocker, F., Johnson, R. L., Starr, E. L., Hall, A. M., Preti, D. M., Slate, S. R., et al. (2019). The time course of processing emotion-laden words during sentence reading: evidence from eye movements. *Acta Psychol.* 192, 1–10. doi: 10.1016/j.actpsy.2018.10.008

Knickerbocker, H. (2014). *Release From Proactive Interference: The Impact of Emotional and Semantic Shifts on Recall Performance* (unpublished doctoral dissertation). University at Albany, State University of New York, Albany, NY (2014).

Knickerbocker, H., and Altarriba, J. (2013). Differential repetition blindness with emotion and emotion-laden word types. *Vis. cogn.* 21, 599–627. doi: 10.1080/13506285.2013.815297

Lench, H. C., Flores, S. A., and Bench, S. W. (2011). Discrete emotions predict changes in cognition, judgment, experience, behavior, and physiology: a meta-analysis of experimental emotion elicitations. *Psychol. Bull.* 137, 834–55. doi: 10.1037/a0024244

Lin, S. Y., Chen, H. C., Chang, T. H., Lee, W. E., and Sung, Y. T. (2019). CLAD: a corpus-derived Chinese Lexical Association Database. *Behav. Res. Methods* 51, 2310–2336. doi: 10.3758/s13428-019-01208-2

Pérez, M., Stadthagen-Gonzalez, H., Guasch, M., Hinojosa, J. A., Fraga, I., Marin, J., et al. (2021). EmoPro – Emotional prototypicality for 1286 Spanish words: relationships with affective and psycholinguistic variables. *Behav. Res. Methods.* 1–19. doi: 10.3758/s13428-020-01519-9

Shiota, M. N., Campos, B., Oveis, C., Hertenstein, M. J., Simon-Thomas, E., and Keltner, D. (2017). Beyond happiness: building a science of discrete positive emotions. *Am. Psychol.* 72, 617–643. doi: 10.1037/a0040456

Sutton, T. M., and Altarriba, J. (2016). Color associations to emotion and emotion-laden words: a collection of norms for stimulus construction and selection. *Behav. Res. Methods* 48, 686–728. doi: 10.3758/s13428-015-0598-8

Wang, X., Shangguan, C., and Lu, J. (2019). Time course of emotion effects during emotion-label and emotion-laden word processing. *Neurosci. Lett*. 699, 1–7. doi: 10.1016/j.neulet.2019.01.028

Wu, C., and Zhang, J. (2019a). Conflict processing is modulated by positive emotion word type in second language: an ERP study. *J. Psycholinguist. Res.* 48, 1203–1216. doi: 10.1007/s10936-019-09653-y

Wu, C., and Zhang, J. (2019b). Emotion word type should be incorporated in affective neurolinguistics: a commentary on Hinojosa, Moreno and Ferré (2019). *Lang. Cogn. Neurosci.* 35, 840–843. doi: 10.1080/23273798.2019.1696979

Wu, C., Zhang, J., and Yuan, Z. (2019). An ERP investigation on the second language and emotion perception: the role of emotion word type. *Int. J. Bilingual Educ. Bilingualism*. doi: 10.1080/13670050.2019.1703895. [Epub ahead of print].

Wu, C., Zhang, J., and Yuan, Z. (2020). Affective picture processing is modulated by emotion word type in masked priming paradigm: an event-related potential study. *J. Cogn. Psychol.* 32, 287–297. doi: 10.1080/20445911.2020.1745816

Wu, C., Zhang, J., and Yuan, Z. J. B. S. (2021). Exploring affective priming effect of emotion-label words and emotion-laden words: an event-related potential study. *Brain Sci.* 11:553. doi: 10.3390/brainsci11050553

Xu, X., and Li, J. (2020). Concreteness/abstractness ratings for two-character Chinese words in MELD-SCH. *PLoS ONE* 15:e0232133. doi: 10.1371/journal.pone.0232133

Yao, Z., Wu, J., Zhang, Y., and Wang, Z. (2017). Norms of valence, arousal, concreteness, familiarity, imageability, and context availability for 1,100 Chinese words. *Behav. Res. Methods* 49, 1374–1385. doi: 10.3758/s13428-016-0793-2

Zhang, J., Teo, T., and Wu, C. (2019a). Emotion words modulate early conflict processing in a flanker task: differentiating emotion-label words and emotion-laden words in second language. *Lang. Speech* 62, 641–651. doi: 10.1177/0023830918807509

Zhang, J., Wu, C., Meng, Y., and Yuan, Z. (2017). Different neural correlates of emotion-label words and emotion-laden words: an ERP study. *Front. Hum. Neurosci.* 11:455. doi: 10.3389/fnhum.2017.00589

Zhang, J., Wu, C., Yuan, Z., and Meng, Y. (2018a). Different early and late processing of emotion-label words and emotion-laden words in second language: an ERP study. *Second Lang. Res.* 36, 399–412. doi: 10.1177/0267658318804850

Zhang, J., Wu, C., Yuan, Z., and Meng, Y. (2019b). Differentiating emotion-label words and emotion-laden words in emotion conflict: an ERP study. *Exp. Brain Res.* 237, 2423–2430. doi: 10.1007/s00221-019-05600-4

Zhang, J., Wu, C., Zhou, T., and Meng, Y. (2018b). Cognate facilitation priming effect is modulated by writing system: evidence from Chinese-English bilinguals. *Int. J. Bilingualism* 23:1367006917749062. doi: 10.1177/1367006917749062

Zhou, Y., and Tse, C.-S. (2020). The taste of emotion: metaphoric association between taste words and emotion/emotion-laden words. *Front. Psychol.* 11:986. doi: 10.3389/fpsyg.2020.00986

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.

frontiers in Communication

ORIGINAL RESEARCH
published: 17 November 2021
doi: 10.3389/fcomm.2021.729751

Pumping Up the Base: Deployment of Strong Emotion and Simple Language in Presidential Nomination Acceptance Speeches

Cynthia Whissell *

Psychology Department, Laurentian University, Sudbury, ON, Canada

Background and Method: This research examines the nomination acceptance speeches of US presidential candidates from Republican and Democratic parties in the post-WWII mass communication era (1948–2020, 38 speeches). Variables studied are the emotional tone of the speeches, their abstractness, their Grade Level, their employment of personal pronouns and their mentions of "America". Speeches were scored with the Dictionary of Affect in Language (a sentiment analysis tool).

Predictions: On the basis of functionalist theories of political discourse, it was predicted that the speeches would have a pleasant and active or celebratory emotional tone. Based on related research that focused on the effects of mass distribution on presidential communications, it was predicted that the speeches would increase in pleasantness, arousal and linguistic simplicity across years.

Results: As predicted, speeches were pleasant and active in tone. Across years, speeches became significantly more arousing, less abstract, simpler, and longer. When individual speeches were divided into five equal portions, a strong significant quadratic trend was observed for pleasantness, which started high at the beginning of a speech, fell in the center, and rose again at the end.

Conclusions: Presidential nomination acceptance speeches are emotionally pleasant and active and linguistically simple (Grade 8 level). Between 1948 and 2020, they remained pleasant, and became more active and simpler. In service of their aim to "pump up the base" individual speeches began on a pleasant, nationalistic and personal note, encompassed duller and more impersonal material in their centers, and became positive again at the end.

Keywords: presidential nomination contests, emotion, language complexity, functionality, mass distribution

OPEN ACCESS

Edited by:
Daniel Broudy,
Okinawa Christian University, Japan

Reviewed by:
Karen Sebold,
University of Arkansas, United States
Matthew Eshbaugh-Soha,
University of North Texas,
United States

Correspondence:
Cynthia Whissell
cwhissell@laurentian.ca

Specialty section:
This article was submitted to
Political Communication and Society,
a section of the journal
Frontiers in Communication

Received: *23 June 2021*
Accepted: *10 September 2021*
Published: *17 November 2021*

Citation:
Whissell C (2021) Pumping Up the Base: Deployment of Strong Emotion and Simple Language in Presidential Nomination Acceptance Speeches. Front. Commun. 6:729751. doi: 10.3389/fcomm.2021.729751

INTRODUCTION

Functionalist theories of rhetoric (Benoit, 1999, Benoit, 2014; Benoit et al., 1998, Benoit et al., 2004) posit that political speeches, including the presidential nomination acceptance speeches of modern candidates, are instrumental in nature. Form follows function and speeches are deliberately structured to promote desired outcomes. The emotional tone of speeches is part of the arsenal deployed with intent to influence an audience and accumulate votes. Nomination acceptance speeches are celebratory, coming as they do at the end of a long selection process; they also serve to unify the party, to provide a climactic moment for the convention, and to outline the platform on which the candidate will run (Benoit, 1999). Attack and defense are two of the main functions of political speeches (Benoit et al., 1998), with acclaim, which is key to acceptance speeches, being a third. Whissell (2012) reported on the thrust and parry nature of interchanges in the Canadian parliament's Question Period: questioners from opposition parties tended to use more aggressive language then responders who spoke for the ruling party in attempts to deflect attacks. Benoit (1999) noted that nomination acceptance speeches tended to focus on acclaim rather than attack, in a ratio of roughly three acclaims to each attack. Such findings lead to the expectation that presidential nomination acceptance speeches will have a positive and active emotional tone that promotes celebration and unity and reflects the acclamatory mood of the speeches.

This research examines presidential nomination acceptance speeches with the help of a sentiment analysis tool, the Dictionary of Affect in Language (Whissell, 2009). Such tools quantify the language employed in communications in terms of its underlying emotional tone (Young and Soroka, 2012). The Dictionary has been applied to the study of presidential inaugural addresses (Whissell and Sigelman, 2001) and presidential radio communications (Sigelman and Whissell, 2002), as well as other types of political communications (Sigelman, 2002; Whissell, 2012). In its present form (Whissell, 2009), the Dictionary provides scores for the pleasantness or positivity of a text, its activation or arousal, and the extent to which the words in the text promote strong visual imagery. Scoring is accomplished on a word by word basis. The presence of extremely pleasant words (e.g., hope, happy, win, wisdom, comfort, and successes) adds to the overall positivity of a speech while the presence of unpleasant words (e.g., sick, alone, argued, attacked, failure, and turmoil) decreases it. A speech containing many active words (e.g., adventure, energetic, discovery, nightmare, weapons, and violently) has a more arousing tone while one containing passive words (e.g., solemn, senior, slowed, ending, ignorance, and humble) is calmer and quieter in tone. Finally, an acceptance speech containing many words that are easily envisioned (e.g., dollars, children, people, guns, banks, and newspaper) would score higher on imagery or concreteness, while one which contained words that do not easily promote mental pictures (e.g., conduct, preservation, opportunity, finally, justify, and dignity) would be abstract. The examples provided here are extreme words from the nomination acceptance speeches studied. Imagery, the third dimension of the Dictionary of Affect, is not, strictly speaking, an affective dimension. It is a cognitive one (Vecchi, 2019). Images associated with words, (and therefore speeches) can make them easier to remember (Marre et al., 2021). When emotions or abstract concepts are being discussed, for example when people describe their feelings or when they debate the meaning of "justice", the words employed tend to be abstract and difficult to image. They do not easily stimulate mental pictures representing their meanings. This does not imply that the meaning of these words is unclear, but rather that the words are associated with non-pictorial concepts.

The first prediction of the research (P1), is that presidential nomination acceptance speeches will have a positive, active and abstract tone. This prediction will be tested by comparing the three scale means from Dictionary of Affect scoring of the speeches to normative values for a large sample of everyday English. It is specifically expected that pleasantness and activation values for presidential nomination acceptance speeches will be above the normative value while imagery will be below it.

Whissell and Sigelman (2001) pointed out that presidential inaugural addresses changed across time, as a function of their expected audience, with speeches becoming simpler and more emotional as the audience grew larger, thanks to developments in mass communication. On this basis, it was predicted (P2) that nomination acceptance speeches would follow the same pattern, becoming more pleasant and active, and also simpler, across time. The prediction of simplicity has two parts: it was expected that speeches would become more concrete, and therefore have higher scores for imagery across time, and it was predicted that the Grade Level (Flesch-Kincaid Grade Level) of the speeches would fall across time. Correlations of yearly mean values with year itself were employed to test this second set of predictions.

Two additional types of measures were added to the research, but these were treated in an exploratory manner and were not associated with specific predictions. The 1996 film comedy "My Fellow Americans" takes its name from a phrase commonly associated with the opening of political addresses, where it does, in fact, frequently appear (Waldeman, 2010). This phrase and its variants ("My fellow Republicans"; "My fellow Democrats") appear repeatedly in presidential nomination acceptance speeches. The pronoun "my", and the word "fellow" join the audience to the speaker. The use of pronouns such as "my" in political speeches was examined by a variety of researchers (e.g., Bull and Fetzer, 2006; Savoy, 2010; Vertommen, 2013; Putri and Kurniawan, 2015). Four additional measures of the research were associated with the frequency of use of individual words: first-person pronouns (e.g., I, we), second-person pronouns (e.g., you, your), third person pronouns (e.g., he, she, and they), and the use of the word "America" in its various forms. The final type of measure addressed within-speech structure for the addresses. Several researchers who analyzed presidential nomination speeches (e.g., Miller and Styles, 1986; Benoit, 1999; Deason and Gonzales, 2012) have examined them as a unit. In this research is was decided to also examine the speeches in five parts or portions, with each portion of a speech containing a roughly equal number of words. If speeches are structured to a purpose, it is likely that their beginning, middle, and end would serve different functions and yield different values for the variables under study.

METHODS

Presidential nomination acceptance speeches were downloaded in March and April, 2021, from the American Presidency Project of the University of California at Santa Barbara. Most, but not all acceptance speeches were given at conventions. The most recent exception is Joe Biden's acceptance speech which was delivered at a convention but remotely, due to COVID-19 pandemic restrictions. Comments by the audience were removed from all speeches. Each speech was then scored by an SPSS program that compared every word in it to the Dictionary of Affect and imported scores for any matched words into a data file. The 38 speeches contained a total of 157,956 words, 93% of which had scores associated with them for pleasantness, activation, and imagery. A simple matching-and-counting analysis for string variables in SPSS was employed to identify the relative frequency of first-person, second-person, and third person personal pronouns, and the use of America-words (America, American, and Americans…). Speeches were divided into five equal parts or portions on the basis of the length of each speech. Speeches were scored using the readability statistics available in Word® under the Spelling and Grammar function in Review. This function produces a Flesch-Kincaid Grade Level, which reflects the grade of readers for whom the text would be appropriately complex (Stockmeyer, 2009). Grade Levels of 9–12 would be suitable for high school students, and Grade Levels of 13–16 for university students. A Grade Level of 18 would suggest that the text was appropriate for graduate students. The Grade Level for this paragraph is 13.6.

Although the Dictionary has a high scoring rate (typically 90%, 93% in this research, with a standard deviation across speeches of 1%), there are still issues associated with its use. One such issue is the possible multiple meanings of single words and another the very high natural variability of language. The word "read" for example, is used in many different ways. It represents several tenses of the same original verb, and it can also serve as a noun ("That was a good read!"). The Dictionary of Affect does not report multiple meanings, as there is only one set of ratings in it associated with the word "read" which is very pleasant, quite active, and very easy to image. It is assumed that each of the raters who originally provided the dimensional scores for the Dictionary was thinking of one possible meaning of "read", but not necessarily of the same meaning. The scores in the Dictionary therefore represent "What people think of when confronted with a word," and do not differentiate between possible multiple meanings. The shortcoming of ambiguous meanings is characteristic of any dictionary-scored system of language analysis. In this research, thousands of words were scored for the speech of every candidate, and matching rates were uniform, so it is unlikely that multiple meanings would differentially affect the speeches of different candidates. As well, a few confusing words in a large sample can hardly skew the result for the whole sample that contains thousands of scored words. Rudkowsky and co-authors (Rudkowsky and Haselmeyer, 2018), refer to dictionary- or lexicon-based methods of sentiment analysis as "bags-of-words" approaches, but they also recognize that these are the dominant form of sentiment analysis. Improvements in sentiment analysis could attempt to address context in order to disambiguate the meanings of words. For this approach to succeed, systems will almost certainly need to be trained in recognizing rhetorical tactics such as sarcasm and irony. Although training (machine learning) is possible, it is labor intensive and likely to involve domain-specific crowd-coding, which embodies human opinions, and is employed to train programs to recognize the meanings of words more specifically.

Natural language is extremely variable in terms of the emotional tone of its words, so that even one sentence might contain words with many different possible emotional impacts. For example, the sentence "I thank you with all my heart for your friendship and your confidence," which occurs in Dewey's 1948 nomination acceptance speech has an overall positive emotional tone (69, where the norm for everyday English is 50). However, along with words that are very pleasant (e.g., heart, 99; friendship, 89) the sentence includes some that are somewhat less pleasant (e.g., confidence, 42). Function words (e.g., and, 39; for, 33) are part of almost every sentence. Dewey's sentence has a standard deviation of 24.3. In absolute terms, this is greater than the standard deviation of all the words in all the speeches studied (21.8). When language is naturally variable, effect sizes associated with analyses at the level of the word will tend to be small. It is true that some single words carry extreme emotional scores, but a normal sentence has a structure that demands it contain words with very different emotional tones. Small effect sizes are neither insignificant nor meaningless. The tone of an entire speech can be changed by the inclusion of a few extreme words in it, but the speech as a whole will remain emotionally variable.

RESULTS

The Overall Character of the Speeches

Mean pleasantness for all speeches, with words as the unit of measurement, was 50.66; mean activation was 51.20, and mean imagery 48.67. These means were compared to the norm for everyday English which was 50 with a t test for one mean. All means were significantly different than 50 (t_{146376} = 11.55, 21.22, and −14.08 respectively, $p < 0.001$). The speeches were pleasant, active, and abstract in comparison to everyday English. Effect sizes were weak with d = 0.03, 0.08, and 0.04. These effect sizes should be interpreted in view of the natural variability of language discussed above. The most notable emotional characteristic of the speeches was their high activation: this is likely tied to their function, which is "pumping up the base" of supporters. The mean grade level of the speeches was 8.56, which was significantly lower than grade 12 (the last grade in secondary school; $t_{37} = -13.22$, $p < 0.001$, $d = 1.61$). The grade level of 8.56 implies that the text of the speeches is appropriate for students half-way through grade 8. Web sites such as *The Readable Blog* (https://readable.com/blog/what-is-the-average-persons-reading-level/) suggest that, in order to reach a maximum audience, materials on the web should be at Grade 8 level. The average reading level of Americans is Grade 8, so the nomination acceptance speeches were accessible to a very wide audience. The first prediction of the study (P1) was confirmed with respect to all variables. Presidential nomination acceptance speeches were pleasant, active, somewhat abstract, and linguistically simple.

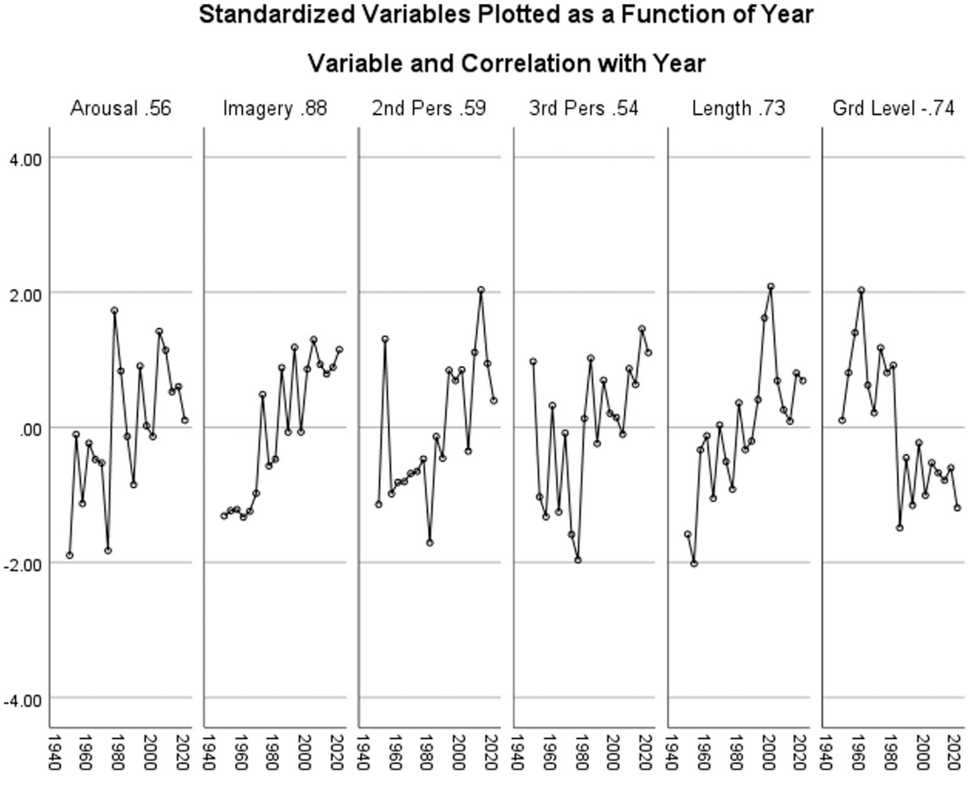

FIGURE 1 | Changes across time for six standardized variables significantly related to year.

Changes in Speeches Across Time

Mean values were calculated for each variable in each year by averaging the data for the two candidates in that year. When these mean values were correlated with year (19 observations), six significant correlations were noted (**Figure 1**). Arousal, imagery, use of first and second person pronouns, and length all increased significantly with year. Grade Level decreased significantly. Correlations for arousal and imagery were in accordance with predictions (P2), but pleasantness showed no change across time. The expectation of greater simplicity across time (also from P2) was supported by the negative correlation of Grade Level with year. Scores significantly related to year were standardized in order to make them comparable, and changes across time are plotted in **Figure 1**.

Differences Among Individual Candidates' Speeches

Means for each candidate's speech for each variable are noted in **Supplementary Appendix SA**. A multivariate analysis of variance was employed to compare the 38 speeches: Grade Level and length in number of words were not included in this analysis because these scores characterized entire speeches. According to a multivariate analysis of variance with word as the unit of observation and speaker as the only factor, there were significant differences among speakers for every variable (multivariate effect, $p < 0.001$, $\eta^2 = 0.002$; univariate effects, $2.83 < F_{37,\ 146{,}339} < 9.96$; $p < 0.001$; $0.001 < \eta^2 < 0.003$). Speech lengths were compared using a goodness of fit Chi Squared statistic which indicated that some speeches were significantly longer than others ($\chi^2_{37} = 21{,}369.07$, $p < 0.001$). **Table 1** reports mean values for the two most recent speakers (Trump and Biden, 2020) along with the mean for the entire group and an estimate of two standard errors. Comparisons between any two speakers equivalent to a two-tailed *t* test with *p* at 0.05 can be made by subtracting one mean from the other and comparing the result to the estimate of two standard errors. When the difference exceeds the estimate, it is significant. For example, the difference between Trump and Biden for pleasantness was 0.47: this did not exceed the value of two standard errors for this variable (0.71) and was therefore not a significant difference. However, the difference between the two speakers for imagery was 1.75, which is greater than the value of two standard errors (1.16), and therefore significant. Trump's acceptance speech was more highly imaged and Biden's was more abstract. Continuing the comparison between these two speakers, there was no difference for activation, second-person pronouns and the employment of America-words, but Biden employed relatively more first-person pronouns and Trump more third-person pronouns. Trump's speech (6,980 words) was more than twice as long as Biden's (3,202). Biden's speech (Grade 5.8) was pitched at a lower Grade Level than Trump's (8.1), which was closer to the group mean Grade Level of 8.6. Comparisons within any column of **Supplementary Appendix SA** can be made in a similar

TABLE 1 | Means for all speech characteristics for the two most recent presidential nomination acceptance addresses: Trump and Biden, 2020.

Characteristic	Candidate Trump	Candidate Biden	Overall Mean (and 2[a]se)
Year	2020	2020	
Party	Republican	Democrat	
Pleasantness	50.42	50.89	50.66 (0.71)
Arousal	51.56	50.85	51.20 (0.71)
Imagery	51.05[a]	49.30	48.67 (1.16)
Use of First Person Pronouns	0.051	0.060[a]	0.063 (0.008)
Use of Second Person Pronouns	0.013	0.010	0.012 (0.004)
Use of Third Person Pronouns	0.027[a]	0.017	0.019 (0.005)
Use of America Words	0.014	0.013	0.011 (0.002)
Length (in Words)	6,980	3,202	4,157
Flesch-Kincaid Grade Level	8.1	5.8	8.6

Note: For pronoun use and use of America words, the data are proportions.
[a]*This mean was significantly higher,* $p < .05$.

manner. Any difference between means greater than the estimate of two standard errors is statistically significant.

Structure of Speeches

Seven mixed analyses of variance with portion of the speech (1st through 5th) as a repeated measure and party (Republican, Democrat) and status (eventual winner of presidency, eventual loser) as between-subjects variables were conducted for the outcome variables of pleasantness, activation, imagery, use of first-person, second-person, and third-person personal pronouns, and use of America-words. The unit of observation was the mean for each variable for each portion of each speech (190 observations per analysis). There were no significant differences of any kind associated with party. The eventual winner factor evinced only one significant main effect for the use of second-person pronouns ($F_{1,34} = 6.45$, $p = 0.016$, $\eta^2 = 0.16$). Eventual winners employed many fewer second-person pronouns than eventual losers (9 per 1,000 words as opposed to 120). There were no significant interactions involving status. There were significant main effects associated with portion of the speech for four variables: pleasantness ($F_{4,136} = 14.57$, $p < 0.001$, $\eta^2 = 0.30$), the use of America-words ($F_{4,136} = 3.30$, $p = 0.013$, $\eta^2 = 0.088$), and the use of first-person ($F_{4,136} = 4.01$, $p = 0.004$, $\eta^2 = 0.106$) and second-person ($F_{4,136} = 10.71$, $p < 0.001$, $\eta^2 = 0.240$) pronouns. In all cases where the main effect for portion was significant, the trend observed was a strong quadratic one ($p < 0.001$, $\eta^2 = 0.679$ for pleasantness; $p < 0.001$, $\eta^2 = 0.268$ for America words; $p < 0.001$, $\eta^2 = 0.356$ for first-person pronouns; $p < 0.001$ and $\eta^2 = 0.393$ for second-person pronouns). There were no other significant effects. The strongest quadratic effect (with an overwhelming $\eta^2 = 0.68$) was associated with pleasantness. The U-shaped quadratic trends pictured in **Figure 2** began with high scores in the first portion (higher pleasantness, more references to America, more first and second-person pronouns), dropped to lower scores for the middle portions, and rose again at the end of the speech. These changes are probably related to the fact that acceptance speeches normally open and close to the raucous acclaim and cheers of the audience.

Some *Post Hoc* Analyses

4 Four *post hoc* analyses were undertaken in order to explore differences among speeches. In the first analysis, the Mahalanobis distance measure was computed to search for multivariate outliers among the speeches in terms of the nine dependent variables: no outliers were identified as all distances had associated p values >0 .01. Next, the popularity of each candidate and their signed difference in popularity from their main competitor in July of the election year, taken from Gallup poll data, were correlated with all the dependent variables. No significant relationships were noted ($p > 0.05$). In his commentary, Mayer (2004) suggested that changes in campaign strategy might be associated with two inflection points: 1972, when the more modern convention system gelled and 1996, when the large gap between the clinching of the nomination and the convention became evident. Dummy variables were created for each of these dates, with years 1972 and later coded as 1 for the first variable and years 1996 and later coded as 1 for the second. Year and the two dummy variables were correlated with the dependent variables. Relationships for the dummy variables were often significant, but they were not stronger than relationships for year, so the dummies did not add any information to the analysis. The correlation coefficients can be viewed in **Supplementary Appendix SB**. In the final analysis, nomination acceptance speeches by incumbent presidents were compared to those made by other candidates. This analysis involved t-tests at the level of individual words ($df > 100{,}000$ in all cases, tests at $p < 0.05$). Incumbents employed significantly less pleasant language (50.42, 50.78, $d = 0.017$), more active language (51.36, 51.12, $d = 0.011$), more highly imaged language (49.20, 48.39, $d = 0.022$), fewer first-person pronouns (0.055, 0.061, $d = 0.025$), fewer second person pronouns (0.010, 0.011, $d = 0.011$) and more third person pronouns (0.019, 0.017, $d = 0.016$). There was no significant difference for America words. As indicated by d values, the effect sizes involved were small. In comparison to incumbents, non-incumbent candidates used more pleasant and less active and imaged language, and the tended to employ more first and second person pronouns and fewer third

person pronouns. Their speeches were therefore more personalized, and more directed at pleasing their audience.

DISCUSSION

Overview of Results

This research examined presidential nomination acceptance speeches by quantifying them in terms of nine variables: their emotional tone (pleasantness, activation), their linguistic simplicity (imagery, Grade Level), their use of personal pronouns (first, second, and third person), their use of forms of the word "America", and their length. Several predictions were made with respect to the overall character of nomination acceptance speeches. All of them were confirmed. Speeches were pleasant and active in tone, and they were also somewhat abstract but not linguistically complex. A second set of predictions addressed changes across time in acceptance speeches. All but one of the predictions in this set were also supported by the data. The speeches became more active and more concrete across time, and their language also became simpler: it had lower Grade Levels. Speeches also grew longer over time.

Exploratory analyses indicated that individual speeches differed in terms of all possible measures. A method was provided (**Table 1**, **Supplementary Appendix SA**) to allow researchers interested in making individual comparisons among speakers to do so. Further exploratory analyses addressed the structure of speeches, looking at how the measures of the study changed across the five portions of each speech (beginning, 2nd fifth, middle, 3rd fifth, and ending). Significant trends were identified for pleasantness, the use of America-words, and the use of first and second person pronouns. In all cases, the strongest significant trend was quadratic, pointing to a strong opening in terms of these four variables, a fall to the centre portions of the speech, and a strong closing. Speeches were all about "you", "me" and "America" in their beginnings, and their emotional tone was pleasant: platforms and issues tended to be discussed in the 2nd, 3rd, and 4th portions of the speeches, and the last portion resounded again with "you", "me", and "America", closing on a positive note (**Figure 1**). Additional *post hoc* analyses examined differences between candidates who were incumbent presidents and non-incumbents. Differences observed suggest that the language of incumbents was characterized by a more reassured and less personalized approach.

Some of the differences reported in this research—for example those associated with differences between incumbents and non-incumbents or those associated with P1—were significant but weak. This can be tracked back to the natural variability of language. Weak differences are still meaningful but they represent a small change in the overall product (the speech). The strongest effect sizes associated with the speeches as a whole were those representing their Grade Level and their activation: the base was being "pumped up" by the nominees' use of active and accessible language. The strongest effect sizes associated with changes across time were for language simplicity, i.e., increased imagery and decreased Grade Level, and for length (speeches became longer). The strongest effect size in the study of speech structure was associated with pleasantness which started high, fell in the middle of the speeches, and rose again at the end. Members of the nominees' base left the convention in an active and positive frame of mind, with the sense that the nominees had been speaking to them personally. It is notable that neither party nor the eventual status of the candidate as winner of the presidency (or not) were associated with strong differences among speeches. As a limitation, it should be noted that this research provides an overview of nomination acceptance speeches, their structure, and their differences, and that much information remains within them that could be studied at a deeper level in future research. Furthermore, changes in the ways in which nominations were made appeared at several times during the period 1948–2020, as

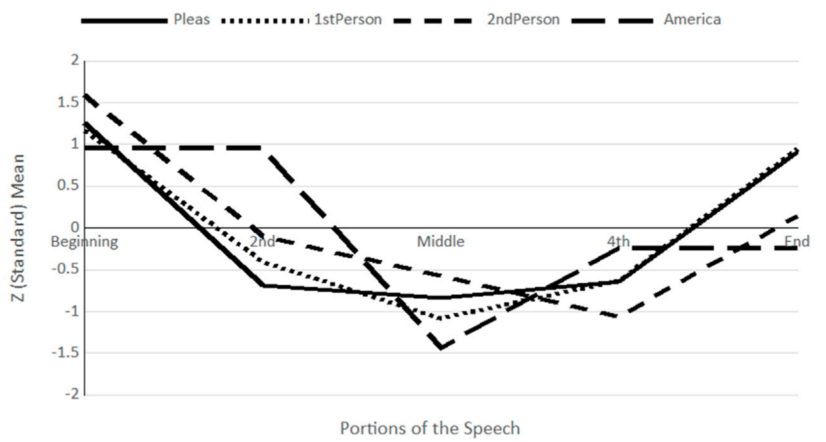

FIGURE 2 | The structure of speeches: changes across five equal portions of speeches for pleasantness and the use of first person and second person pronouns and America-words. Note: to allow for comparability, all means were plotted as standard (z) scores.

did changes in the style and extent of media coverage. Both of these might have influenced results.

Functionality of Political Speeches and Their Mass Distribution

The predictions of this research relied on the presence of two main forces that were assumed to impact on political discourse. The first force was the functionality of such discourse and the second its mass distribution. Speeches are shaped in the way they are because their structure serves a particular function, and they are framed in more or less complex or emotional terms because it is expected that many, many Americans will engage with them, in one form or another. In their work on *The Ubiquitous Presidency*, Scacco and Coe (2021) brought both of these forces into focus simultaneously:

> The [president's] speech will draw a massive audience, one of the biggest of the year. Those who don't encounter the speech live will see a television news recap later in the evening or read about it in tomorrow's paper . . . Every word that passes the president's lips has been carefully crafted for this occasion. If he can get the message just right, the thinking goes, the public will understand his perspective and help him achieve his political aims. . . .

[Moments are] produced by presidents, reported and repurposed by journalists, studied by scholars, witnessed and talked about by citizens. In combination, these moments form a mythology that lives on in books and on websites, in high school and college classes, in popular portrayals of the presidency, and in the public imagination. (p. 1).

The authors go on to make the point that the presidency has morphed in recent years, and has gone from being a mainly rhetorical presidency to being a ubiquitous one, where news is managed in much more complex ways than it had been in the past. One might consider nomination acceptance speeches to be less relevant and impactful than, for example, inaugural addresses or messages on the State of the Union. They are, however, part of the total package of carefully managed political language that focuses relentlessly on its goal of access to power. The names of speakers are identifiers: it is suggested that each of them represents his party's drive for success in any given year, as well as representing themselves. The speaker does impact the speech—Reagan and Trump's speeches, for example, echo their distinctive uses of language—however this is accomplished within a solid understanding of the purpose of the speech (to lead to an electoral win) and the nature of its audience (a very wide one).

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

CW downloaded and analysed the data. She also wrote the article.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcomm.2021.729751/full#supplementary-material

REFERENCES

Benoit, W. L. (2014). A Functional Analysis of 2008 and 2012 Presidential Nomination Acceptance Speeches. *Speaker. Gavel.* 51, 50–59.

Benoit, W. L. (1999). Acclaiming, Attacking, and Defending in Presidential Nominating Acceptance Addresses, 1960-1996. *Q. J. Speech.* 85, 247–267. doi:10.1080/00335639909384260

Benoit, W. L., Blaney, J. R., and Pier, P, M. (1998). *Campaign '96: A Functional Analysis of Acclaiming, Attacking, and Defending.* Westport, Connecticut: Praeger.

Benoit, W. L., Stein, K. A., and Hansen, G. J. (2004). How Newspapers Cover Presidential Nomination Acceptance Addresses. *Newspaper Res. J.* 25, 83–89. doi:10.1177/073953290402500308

Bull, P., and Fetzer, A. (2006). Who Are We and Who Are You? the Strategic Use of Forms of Address in Political Interviews. *Text Talk.* 26, 3–37. doi:10.1515/text.2006.002

Carlton, W. G. (1957). The Revolution in the Presidential Nominating Convention. *Pol. Sci. Quart.* 72, 224–240. doi:10.2307/2145774

Deason, G., and Gonzales, M. H. (2012). Moral Politics in the 2008 Presidential Convention Acceptance Speeches. *Basic Appl. Soc. Psychol.* 34, 254–268. doi:10.1080/01973533.2012.674450

Marre, Q., Huet, N., and Labeye, E. (2021). Embodied Mental Imagery Improves Memory. *Q. J. Exp. Psychol.* 1, 174702182110092. doi:10.1177/17470218211009227

Mayer, W. G. (2004). From the End of the Nomination Contest to the Start of the National Conventions: Preliminary Thoughts on a New Period in Presidential Campaign Politics. *The Forum.* 2 (2), 0000102202154088841032. doi:10.2202/1540-8884.1032

Miller, N. L., and Stiles, W. B. (1986). Verbal Familiarity in American Presidential Nomination Acceptance Speeches and Inaugural Addresses (1920-1981). *Soc. Psychol. Q.* 49, 72–81. doi:10.2307/2786858

Putri, N. A., and Kurniawan, E. (2015). Person Deixis in USA Presidential Camp Aign Speeches. *English Review. J. English Educ.* 3, 190–200.

Rudkowsky, E., Haselmayer, M., Wastian, M., Jenny, M., Emrich, Š., and Sedlmair, M. (2018). More Than Bags of Words: Sentiment Analysis with Word Embeddings. *Commun. Methods Measures.* 12, 140–157. doi:10.1080/19312458.2018.1455817

Savoy, J. (2010). Lexical Analysis of US Political Speeches. *J. Quantitative Linguistics.* 17, 123–141. doi:10.1080/09296171003643205

Scacco, J. M., and Coe, K. (2021). *The Ubiquitous Presidency: Presidential Communication and Digital Democracy in Tumultuous Times.* Oxford: Oxford University Press.

Sigelman, L. (2002). Two Reagans? Genre Imperatives, Ghostwriters, and Presidential Personality Profiling. *Polit. Psychol.* 23, 839–851. doi:10.1111/0162-895x.00310

Sigelman, L., and Whissell, C. (2002). "The Great Communicator" and "The Great Talker" on the Radio: Projecting Presidential Personas. *Presidential Stud. Q.* 32, 137–146. doi:10.1111/j.1741-5705.2002.tb00017.x

Stockmeyer, N. O. (2009). Using Microsoft Word's Readability Program. *Mich. Bar J.* 88, 46–47.

Valley, D. B. (1972). A History and Analysis of Democratic Presidential Nomination Acceptance Speeches to 1968. Urbana-Champaign: University of Illinois. PhD Dissertation. Available at: https://www.proquest.com/docview/302691355?pq-origsite=gscholar&fromopenview=true.

Vecchi, T. (2019). Imagery and Imagination in Psychological Science. *IMGJournal* 1, 314–317. Available at: http://www.img-network.it/wp-content/uploads/2020/01/IMGjournal-ISSUE01-a24-Vecchi.pdf.

Vertommen, B. (2013). The Strategic Value of Pronominal Choice. *Prag* 23, 361–383. doi:10.1075/prag.23.2.07ver

Waldeman, M. (2010). *My Fellow Americans: The Most Important Speeches of America's Presidents from George Washington to Barack Obama*. Naperville, Illinois: Sourcebooks.

Whissell, C., and Sigelman, L. (2001). The Times and the Man as Predictors of Emotion and Style in the Inaugural Addresses of U.S. Presidents. *Comput. Humanities.* 35, 255–272. doi:10.1023/A:1017569003556

Whissell, C. (2012). The Emotionality and Complexity of Public Political Language in Canada's Question Period. *Ells* 2, 68–76. doi:10.5539/ells.v2n4p68

Whissell, C. (2009). Using the Revised Dictionary of Affect in Language to Quantify the Emotional Undertones of Samples of Natural Language. *Psychol. Rep.* 105, 509–521. doi:10.2466/pr0.105.2.509-521

Young, L., and Soroka, S. (2012). Affective News: the Automated Coding of Sentiment in Political Texts. *Politic. Commun.* 29, 205–231. doi:10.1080/10584609.2012.671234

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.

frontiers in Communication

ORIGINAL RESEARCH
published: 02 February 2022
doi: 10.3389/fcomm.2021.758173

Cognitive and Affective Aspects of Verbal Humor: A Visual-World Eye-Tracking Study

*Laura Israel, Lars Konieczny and Evelyn C. Ferstl**

Center for Cognitive Science, Institute of Psychology, University of Freiburg, Freiburg, Germany

Many theories of verbal humour postulate that the funniness of jokes is caused by an incongruency in the punchline whose resolution yields a feeling of mirth. While there are studies testing the prediction that this situation model updating leads to increases in processing costs, there are few studies directly assessing the time course of when the alternative situation models are entertained. In a visual world paradigm, stories were presented auditorily and displays were presented illustrating either the situation implied by the context or the final interpretation after the punchline. Eye movement data confirmed the switch from the initial to the final interpretation for jokes as well as for non-funny control stories that also required a situation model revision. In addition to these effects of the cognitive revision requirements, the pupil dilations were sensitive to the affective component of joke comprehension. These results are discussed in light of incongruency theories of verbal humour.

Keywords: verbal humor, joke comprehension, eye-tracking, visual world, pupil dilation, language, emotion, incongruity

Edited by:
Francesca M. M. Citron, Lancaster University, United Kingdom

Reviewed by:
Jana Lüdtke, Freie Universität Berlin, Germany
Alix Seigneuric, Université Sorbonne Paris Nord, France
Paolo Canal, University Institute of Higher Studies in Pavia, Italy

****Correspondence:***
Evelyn C. Ferstl
evelyn.ferstl@cognition.uni-freiburg.de

Specialty section:
This article was submitted to Language Sciences, a section of the journal Frontiers in Communication

Received: *13 August 2021*
Accepted: *27 December 2021*
Published: *02 February 2022*

Citation:
Israel L, Konieczny L and Ferstl EC (2022) Cognitive and Affective Aspects of Verbal Humor: A Visual-World Eye-Tracking Study. Front. Commun. 6:758173. doi: 10.3389/fcomm.2021.758173

1 INTRODUCTION

Humour is ubiquitous in our social interactions. It is considered relevant for the choice of partners and friends (Bressler and Balshine, 2006), it helps to strengthen in-group cohesion, and it can be an effective tool for relieving tension and reducing stress (Kimata, 2004; Buchowksi et al., 2007).

Psycholinguistic humour research is concerned with the interaction between two aspects of verbal humour: an affective reaction (i.e., the feeling of mirth or enjoyment, possibly shown in smiling or laughing), and a cognitive element, reflecting the correct linguistic interpretation of the intended meaning of the humourous utterances. Goel and Dolan (2001) referred to this distinction as the hot and cold aspects of joke comprehension. In the following we refer to these processes as joke appreciation and joke comprehension, respectively.

The goal of the present study is to use a visual world paradigm to study both aspects. The viewing patterns for pictures illustrating the gist of jokes and appropriate control texts are expected to provide information about the interpretations entertained during listening to the stories. In addition, pupil size measures will be shown to be sensitive to the affective aspects of joke comprehension. Before describing the study in more detail, we will summarize previous psycholinguistic research on verbal humour.

Psycholinguistic research often studies the comprehension of verbal jokes. A prototypical joke, also referred to as canned joke (Martin, 2007) or garden-path joke (Mayerhofer, 2015), is a short, funny story that offers different possibilities of interpretation. After a setting activating a certain situational context, the punchline at the end of the story contains a surprising incongruity whose resolution elicits a feeling of mirth (Dynel, 2009). Incongruity as a crucial prerequisite for humour processing was first described by Kant (1790/1977) and later adapted in incongruity theories of humour (Suls, 1972; Raskin, 1979; Attardo and Raskin, 1991; Giora, 1991; Attardo, 1997; Coulson, 2006). The models assume a two-stage process that can be demonstrated with the following example by Raskin (1979, p. 332):

> "Is the doctor at home?" the patient asked in his bronchial whisper. "No" the doctor's young and pretty wife whispered in reply. "Come on right in".

In the context sentence of this joke the lexical items doctor, patient and bronchial indicate the situation to be a doctor's visit and hence activate a matching narrative scheme or script (also: situation model, van Dijk and Kintsch, 1983; Ferstl and Kintsch, 1999; or frame, Coulson, 2006). This representation of experience—based world knowledge is used to make assumptions about the text continuation. However, the punchline is not consistent with a doctor's visit; the wife's invitation, a surprising violation of the expectations, induces a switch to the correct narrative script—a love affair. Incongruity theories postulate that the funniness of the joke is elicited by the overlap of the two conflicting scripts as well as by the strength of their semantic opposition.

Similar to more general theories of figurative language processing (e.g., Standard Pragmatic View, Grice, 1975; Searle, 1979; Graded Salience Hypothesis, Giora, 2002) incongruity models assume serial processing for the detection and resolution of incongruities. Therefore, they predict that the resolution of an incongruity is more time—consuming and leads to higher cognitive costs than the processing of texts which do not induce situation model updating or frame shifting. In line with this prediction, several studies show a processing disadvantage for jokes over non-funny texts. In a self-paced reading study Coulson and Kutas (1998) presented jokes and matching control texts that differed only in regard to their last word, e.g.,

> I asked the women at the party if she remembered me from last year and she said she never forgets a dress/face/name.

The last word *dress* in this joke triggered an incongruity and a revision, compared to a more expected word, such as *face*. Control texts ended in a similarly unexpected but contextually compatible word, such as *name*. Reading times for the sentence endings were longer in jokes than in the control texts. These results were replicated using eye tracking (Coulson, 2006). In neuroscientific studies, an enhanced N400 component, an indicator for contextual integration difficulties, has been reported for funny texts compared to control texts (Coulson and Kutas, 2001; Coulson and Lovett, 2004; Coulson and Williams, 2005; Mayerhofer and Schacht, 2015). In addition, late integration processes, as indicated by components such as the P600 or the LPC, follow the incongruency detection, but are completed within about 1,500 ms (Canal et al., 2019). Interestingly, there is also evidence that humour facilitates processing. Memory for funny contents or jokes is better than for non-funny texts—a finding that is labelled humour effect (Schmidt, 2002; Strick et al., 2010). In a study comparing jokes to non-funny proverbs, Mitchell et al., 2010 confirmed this memory advantage and reported shorter reading times for the joke punchlines compared to the non-funny proverb endings.

The divergent results of these studies might in part be due to confounds of the funniness of the materials with their revision requirements. In the Coulson and Kutas (1998) study, only the joke ending *dress* requires a frame shift. Both the expected ending *face* and the control ending *name* are aspects of the identity of the person. In the Mitchell et al. (2010) study, on the other hand, inspection of the materials confirms that the proverbs also require processing of non-literal language, and are more difficult than straightforward control texts.

Based on these observations, Siebörger (2006, see also Hunger et al., 2009; Volkmann et al., 2008) argued that the relative contribution of incongruity resolution to joke appreciation requires disentangling the need for revision from the funniness of the stories. Thus, two appropriate control conditions are warranted: First, it is necessary to use identical punchlines to minimize lexical and sentence level differences. Second, and more importantly, the linguistic demands of incongruity resolution need to be compared in funny and non-funny stories. The hypothesis was that incongruity resolution alone is necessary, but not sufficient for making a joke funny, and that comprehension difficulty increases with the cognitive requirements, but does not depend on the emotional content.

To test this hypothesis, jokes and corresponding revision stories were written (Siebörger, 2006). Example materials are shown in **Table 1**. Each text consists of one or two context sentences and a target sentence (punchline). The stories are short dialogues of two protagonists. The punchlines of the jokes and revision texts both contain a sudden twist that causes an incongruity and calls for a revision of the story. The incongruity may depend on a homonym or on a situational shift. That jokes and revision stories differed with respect to their funniness was confirmed in several rating studies (see Siebörger, 2006). Control texts were derived from each joke and each revision story by writing a context sentence that already activated the situation model required for correctly interpreting the punchline. Thus, a straightforward, coherent condition was created with the identical punchline sentences as the experimental stories (jokes or revision texts). This design allowed us to evaluate the effects of linguistic revision independent of word- and sentence-level features of the punchlines. The results of a self-paced reading study included better question answering performance for jokes compared to revision stories and shorter reading times for the punchline of jokes.

In a recent a study using eye tracking during reading we replicated and extended these findings (Ferstl et al., 2017). The more fine-grained temporal analysis confirmed the general

TABLE 1 | Example materials for the joke and revision texts and their matching control texts used in the present study as well as incoherent control texts used by Siebörger (2006) and Ferstl et al. (2017). The punchlines are identical for the experimental stories and their matched control texts. Asterisks denote the revision points.

	Joke	Joke control text
Context	Herbie to his pal Mark: "Why have all these women stopped chasing after you lately?"	Herbie to his pal Mark: "Great that you've managed to stay within the law lately!"
Punchline	Mark: "Well, life as a pickpocket * turned out to be too risky in the long run."	
	Revision Text	**Revision Control Text**
Context	The chef to his wife: "Oh, too bad—the soup is too salty!"	Mr. Miller to his wife: "You oversalted our dinner again!"
Punchline	The wife: "I'm sorry *, I'll try to do better next time!"	

facilitation for jokes, but showed the effects to be due to reading strategies. While first-pass reading times were similar across conditions, total reading times were longer and more regressions were made for non-funny texts compared to jokes. These findings were more pronounced in men than in women, and when the instruction required a meta-cognitive evaluation of the revision requirement, compared to an affective funniness rating.

To explain these effects, we proposed that the appreciation of funniness works as an instantaneous, self-generated feedback about the correctness of the comprehension process (Ferstl et al., 2017). When a the reader laughs at the punchline of a joke it becomes directly apparent that they comprehended the text correctly. The lack of such an intrinsic feedback in non-funny texts might induce rereading strategies, in particular in an experimental setting that promotes a focus on accurate task performance.

While the study was intended to disentangle affective and cognitive components of joke comprehension, both task instructions and the reading format might have induced more deliberate strategies than expected. In addition, although eye tracking during reading is a very sensitive method that provides information about the time course of processing, it gives only indirect evidence about the content of the interpretations entertained at any given moment. Usually, increases in reading times or the occurrence of regressions are taken as evidence for increased processing costs, which in turn are interpreted as reflecting situation model updating.

One important paradigm to assess the contents of the situation model directly is sentence-picture matching (e.g., Zwaan et al., 2002). After reading or listening to a short text, one or several pictures are presented. The comprehender's task is to evaluate whether the picture matches the meaning of the presented text or to select the best illustration among several pictures. Errors and reaction times show whether one interpretation is more easily or more quickly accessible than another. Recently, similar tasks have been adopted for visual world paradigms (Huettig et al., 2011; Berends et al., 2015; Salverda and Tanenhaus, 2018). In this method participant's eye movements are recorded while they look at a visual scene and simultaneously hear a spoken utterance. The spoken expression is related to one or more objects in the visual display: To study ambiguity resolution, the visual scene contains pictures that are compatible with either alternative interpretation (e.g., Tanenhaus et al., 1995). The question of interest is when and for how long the listeners view the respective images. While this paradigm has been successfully applied to study lexical access or syntactic ambiguity resolution (see Salverda and Tanenhaus, 2018, for review), few studies have targeted aspects of higher level comprehension (e.g., Pyykkönen et al., 2010; Pexman et al., 2011) and none, to our knowledge, the comprehension of verbal humour.

The interrelation between visual attention and language processing has been proven in several studies. Cooper (1974) was the first to show that eye movements focus on objects of a visual scene that directly or semantically refer to a simultaneously heard verbal utterance. When hearing the word *boat* subjects were more likely to fixate the picture of a boat (resp., lake) than that of other objects. The number of fixations increases in proportion to the semantic similarity between the visual object and the heard utterance (Huettig and Altmann, 2005). Furthermore, eye movements can even be modulated by words that are only anticipated by the listener and not explicitly stated (Altmann and Kamide, 2007). Because of its very good temporal resolution, depicting language-related effects on oculomotor control within 80–100 ms (Altmann, 2011), and the facilitation of measuring unconscious processes, the visual world paradigm is well suited to gain insight into the temporal characteristics of language processing.

The goal of the present study was to take advantage of this paradigm to describe the viewing patterns for pictures illustrating the alternative situation models elicited by the context and the punchline, respectively. Jokes and revision texts as well as their matching coherent control stories from the material of Siebörger (2006; Ferstl et al., 2017) were used. For each joke and each revision story three scenes were constructed: A target depicting the intended, correct meaning of the story, a competitor that showed the situation implied by the context, and an unrelated distractor. Examples are presented in **Figure 1**.

The images are plausible illustrations of the situations implied by the texts, although other possible scenes might also be associated with the text materials (e.g., a car thief, rather than a pickpocket, as an illustration of a criminal in the example in **Figure 1**).

After the auditory stimulus presentation, participants performed a picture matching task to assess comprehension accuracy and they rated the funniness of the stories. The funniness ratings are expected to be higher for jokes than for all other conditions—which serves as a manipulation check for the stimulus material.

The eye tracking data provided two sources of information. First, the durations of the viewing times (dwell times) of the three pictures were analysed during listening. The viewing times provide evidence for whether and when the depicted situation models are entertained. Even though the

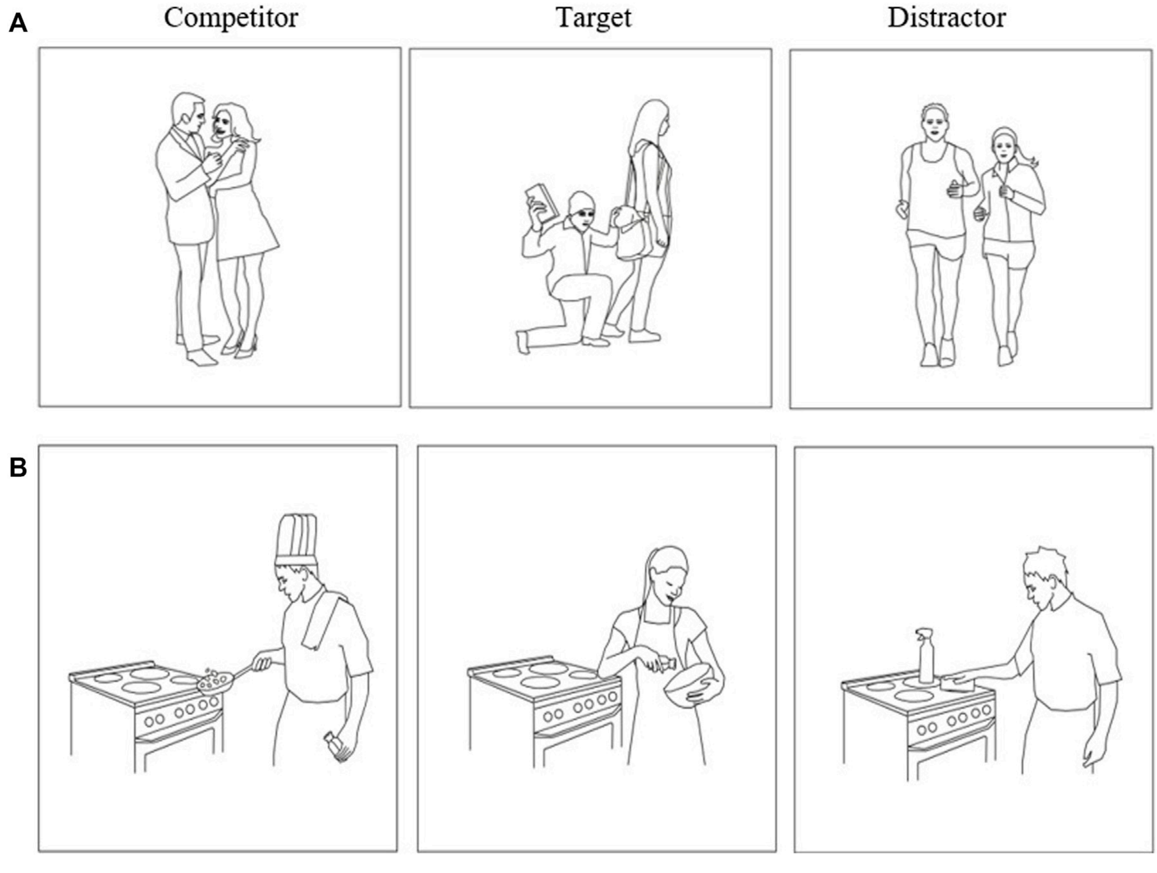

FIGURE 1 | Examples for the visual displays. Shown are the competitor, target and distractor pictures for the joke **(A)** und revision story **(B)** presented in **Table 1**.

present study analyses eye-tracking patterns during joke comprehension in a rather exploratory way, some assumptions were derived from theory and the experimental design: The viewing times were predicted to reflect the current interpretation induced by the stories. Since in the control stories the context sentences already introduced the correct situation model, participants were expected to look more at the target than at the competitor throughout the entire trial. In contrast, in both joke and revision stories, the competitor pictures were expected to be viewed more during the context sentences, followed by a switch to the target only after the incongruency becomes apparent in the punchline. If the previously reported facilitation effect for jokes over the structurally identical revision stories (Ferstl et al., 2017) had been due to meta-cognitive reading strategies only, no differences between jokes and revision stories are predicted. If, on the other hand, the facilitation was due to immediate joke appreciation, the switch to the intended target picture should be faster or more pronounced for jokes than for revision stories.

The second source of information was the pupil size as a measure of physiological arousal (Hess and Polt, 1960; see Sirois and Bisson, 2014, for review). The pupil diameter varies with cognitive load and processing costs (Hyönä et al., 2007; Engelhardt et al., 2010; Wong et al., 2016). More importantly, it also increases with exposure to emotional stimuli (e.g., facial expressions, or sounds of crying or laughing children; Partala and Surakka, 2003), and it has been a useful measure for emotional reactions in studies on social-affective processes (e.g., Prehn et al., 2013). In a study on joke comprehension, Mayerhofer and Schacht (2015) found increases in pupil sizes that varied with the perceived funniness of their text materials. The influence of emotions on pupil size has been explained by the association of pupil dilation with cardiac acceleration (Bradley et al., 2001) and the sympathetic nervous system (Bradley et al., 2008).

As the pupil dilation varies with cognitive load, we assumed that larger pupil sizes would occur in the ambiguous joke and revision conditions compared to the coherent control stories. The crucial question was whether the emotional arousal elicited by jokes, compared to non-funny stories, would also be reflected in pupil dilation. In that case, joke appreciation was expected to lead to an additional increase in pupil size for jokes compared to non-funny revision stories.

2 MATERIALS AND METHODS

Reproducible scripts, open data and open materials (including all auditory as well as all visual stimuli) are provided via our OSF repository at https://osf.io/qk95j/.

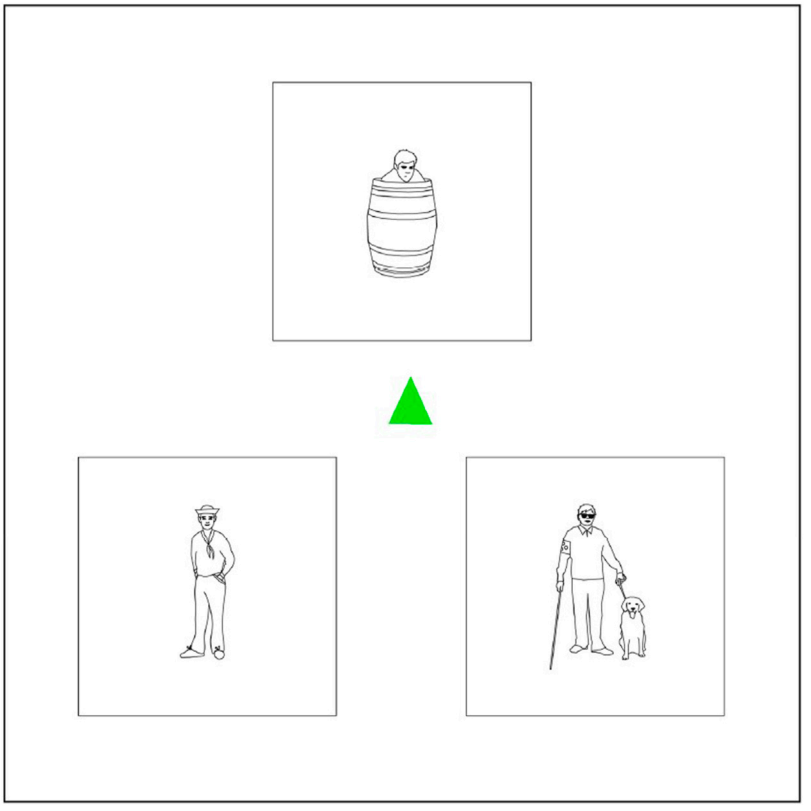

FIGURE 2 | Response displays for the picture matching task.

2.1 Participants

Thirty-one individuals participated in the study. All of the participants were native German speakers and reported to have normal or corrected-to-normal vision. Most of them were students at the University of Freiburg. The data of five participants were excluded. Two subjects had to be eliminated from the sample due to technical problems. Three participants were excluded because of their poor performance in one of the behavioural tasks of the experiment. The final sample used in the analysis of the eye-tracking data consisted of 26 participants (13 women, 13 men). Their mean age was 22.8 years (range: 19-29). Due to the repeated-measures design of the study, where each participant viewed 32 stories, 832 observations were obtained for the evaluation of the behavioural data (funniness ratings and picture matching accuracy).

2.2 Materials

For the experiment 16 jokes and 16 revision stories were selected from the material of Siebörger (2006). The stories were short German dialogues between two characters. All 64 stories had an identical structure, consisting of a context and a target sentence. In the jokes the last sentence (in the following referred to as "punchline") contained information that required a revision of the text and was funny. The revision stories also required a revision, but were not funny. Siebörger (2006) conducted three extensive pretests to select jokes and revision stories for further experiments. Different groups of approximately 32 student participants rated the texts according to their familiarity, distinguished revision and joke stories in an explicit choice, and evaluated whether the text required a situation model revision or not, and how funny they found the texts. 64 jokes and revision texts were selected so that they differed in funniness, could be assigned to one or the other text category, and were unfamiliar to more than 75% of the participants. For the present experiment, a subset of 16 jokes and 16 revision stories was selected for which straightforward illustrations were possible. The two sets were matched for length.

For each experimental joke and revision text a straightforward, coherent control story had been constructed by changing the context sentence and keeping the punchline identical.

For the implementation of the visual world experiment all texts were recorded with the text-to-speech function of the Mac operating system (Version 10.10.3) and adjusted to a natural speaking rate. The automatic voice was chosen to prevent prosodic cues about the story type. All sentences and texts were recorded using the same voice. Audio sequences had an average length of 10,091 ms (range, 7,861–1,174 ms; SD = 999.3 ms). There were no significant differences in length between the text categories. The context sentences and punchlines were recorded separately and connected within the experimental software.

For an analysis of the time course of processing, the individual revision point of each story was defined as the moment at which the incongruity can be perceived and subsequently, the situation model needs to be revised. To determine the revision point for each experimental text, two independent raters were instructed about the structure of the joke and revision stories and were explicitly asked to mark the point where they discovered the incongruity. The revision point usually occurred towards the end of the stories, at about 760 ms before the ending (range 0–4,981 ms). There was no difference between jokes and revision stories with respect to the temporal position of the revision point. The revision point of each experimental story was carried over to its matching control text. The revision points are marked in the example texts in **Table 1**.

For each joke and revision story three black-and-white line drawings were constructed with the Adobe Illustrator (version CS4; Adobe Inc., 2019). The graphics depicted the salient meaning implied by the context (competitor picture), the correct interpretation of the story obtained after the revision process (target picture), as well as an unrelated scene (distractor picture). Examples of the picture sets for joke and revision texts are presented in **Figure 1**. To avoid obvious visual differences between the three pictures, the number of individuals and objects as well as their spatial configuration in the scene were kept similar throughout each set of pictures. The straightforward control stories were presented with the same picture sets as their experimental counterparts. Each scene was presented in a 13 × 13 cm thin black frame. On the visual display the three frames were arranged in a triangular order (**Figure 2**). For the analysis of the eye movements the display was divided into three even interest areas (IAs)—one for each of the presented frames.

To counterbalance the materials, six lists of 32 stimuli were constructed. Each list contained eight experimental jokes, eight experimental revision texts, eight joke control texts, and eight revision control texts. As a result, each punchline sentence was included exactly once in each list. The positions of the pictures (target, competitor, distractor) on the visual display were balanced over the six experimental lists using a Latin square design. The order of the trials within each experimental list was pseudo-randomised with the constraint of not more than two successive trials of the same condition.

2.3 Apparatus and Calibration

The implementation of the experiment and the response recording were conducted with the software package Experiment Builder [version 1.10.165; Experiment Builder (Computer software), 2014]. Eye movements of the dominant eye, assessed in a short test after arrival of the subject, were recorded using the Eye-Link 1,000 eye-tracking system (www.sr-research.com). During the experiment, the participant's head was stabilized with a headrest, about 60 cm away from the computer screen. The eye-tracker was calibrated at the beginning of each block of 16 trials using a nine-point grid. Drift-correction was performed prior to each trial: a black dot was presented in the centre of the screen and stimulus presentation started only after fixation of the dot.

2.4 Procedure

After their arrival participants were randomly assigned to one of the six experimental lists. Written instructions were then presented on the computer screen to explain the task and the calibration procedure. A short training session with four trials followed. For this session pictures from the material of Volkmann, Siebörger and Ferstl (2008) were used. These drawings were more detailed and less balanced than the pictures used in the current study, but were similar in form and content to the experimental materials.

Each trial consisted of an eye-tracking phase followed by two behavioural tasks. In the eye-tracking phase participants first viewed the set of three pictures for 6 seconds. Making the visual displays available before the onset of the auditory presentations ensures that the eye movements reflect the integration of the subsequent linguistic input with the pictures, rather than processes related to visual object recognition and scene interpretation (see Hüttig et al., 2011). The story was then presented over headphones while participants continuously viewed the picture sets. Subsequently, subjects looked at the visual stimuli for another 4 seconds without auditory input. The end of the eye-tracking phase was indicated by a short acoustic signal. At the same time, the blue triangular cursor symbol in the centre of the screen changed its colour to green. Subjects could now move the cursor symbol with their mouse and click on the picture they considered the best match for the meaning of the story. After the picture selection participants were asked to rate the funniness of the story on a nine-point scale ranging from not funny to very funny (cf. Ferstl et al., 2017). The instructions stressed that the participants should use their own intuitive, personal criteria for the ratings, and that they should try to use the full range of values on the scale.

Participants were allowed to take a short break after half of the experiment (16 trials) and they completed the session in about 35–40 min.

2.5 Design and Data Analysis

Statistical analysis was performed using the software R (version 4.1.1; R Core Team, 2017). Mixed-effects regression models were calculated using the lmer und glmer function from the lme4 package (version 1.1-27.1, Bates et al., 2014). A problem in the statistical evaluation of visual world data is that not all observations are independent, because of the multilevel sampling scheme of the paradigm as well as the properties of eye-movements in general (Barr, 2008). To account for these non-independencies, as proposed by Barr (2008), mixed-effects models with random-effects corresponding to the different clusters in the sampling design were used. To increase the interpretability of the very complex models, main effects and interactions were determined using the Anova function from the car package (Fox and Weisberg, 2011). The significance level α was set to $p = 0.05$ (95%). Reported model estimates were calculated using the emmeans function from the lsmeans package (version 2.30-0, Lenth, 2014). For the analysis of the eye-movement data, only correctly answered trials were used to prevent bias due to the misunderstanding of the stories.

To analyse the temporal evolvement of the eye movements two interest periods (IPs) were defined. Each time bin had a length of 1,500 ms. The IPs were time-locked at the story's individual revision point. They covered 3,500 ms, starting 1,500 ms before and ending 2000 ms after the revision point (see Materials). The first IP (before) covered the immediate 1,500 milliseconds before and up to the revision point, while the second IP (after) started 500 ms after the revision point and extended to 2000 ms after the revision point. The first IP was used to establish the baseline, while the second IP was used to detect any effects of the design manipulations. The window duration of 1,500 ms was chosen based on results from Canal et al. (2019) who reported a similar time frame for the completion of joke processing. We excluded the first 500 ms after the revision point to allow for comprehension and integration of the crucial information into the text representation, and to make sure the interest period were clearly separated. The dependent variables included the subjects' performance in the picture matching task (accuracy in percent), and the funniness ratings (on a scale from one to 9). For the analysis of the eye movements, the dwell times (in ms) for each of the three interest areas (IAs) were collected for each IP. In addition, the mean pupil size (measured in arbitrary units; Eyelink Data viewer User's Manual, 2011) was assessed for each IA and each IP. As we expected large interindividual differences in pupil size (e.g. Winn et al., 1994), we z-standardized the pupil size on subject level to remove this random variance from the variable.

For the picture matching task and the funniness rating, the independent variables text category (joke story/revision story) and item type (experimental item/control item) were varied within participants. For the statistical evaluation of the eye tracking data (dwell times and pupil size), IA (target/competitor/distractor) and IP (before vs. after) were added as further within-subject variables. Main and interaction effects of all these variables were added to the models.

In addition, trial (serial position) was added as a covariate to all models.

For the modelling of dwell times and pupil sizes, two funniness covariates were added. First, the group-mean centred (on subject level) funniness ratings were added to the model, coding how funny each story was perceived by each participant in relation to their own average funniness rating. In the following, we will refer to this variable simply as funniness. In addition, we added the mean funniness rating of each participant over all stories as covariate to the models. The variable hence encodes how reactive the respective participant was to the stories in general and how strongly the participant perceived all stories to be humorous. We therefore will refer to this variable as sense of humour. Again, we assume that sense of humour could influence eye-related measures like the dwell times. Besides the two fixed covariates for funniness and sense of humour, a random slope for the funniness variable was introduced for each subject, controlling for possible individual difference in participants' reactions to the stories.

To evaluate whether the performance in the picture-matching task had an effect on eye-tracking variables, the subject's mean performance (percent correct) was introduced as an additional covariate. For the analysis of the pupil sizes, the same fixed effects predictors were used.

The random effects for both models included terms for subject and item. To account for the variability between the stories, a variable that encodes the item and the respective experimental list was added as random factor. Item was coded to match every experimental text (revision or joke) to its associated control text. The experimental list, furthermore, determines the configuration of interest areas on the screen.

The random slopes differed for the analyses of dwell times and pupil sizes. For the dwell times, only IA was used. Because we included fixations on all three relevant interest areas in our models, summed dwell times in each interest period would very often max out to the length of that IP. In addition to adding IA as a fixed effect predictor to our models, we therefore also added the random slope of IA to the dwell time model, to account for individual differences (random effect of subject) or item-baseddifferences (random effect item_and_list) in the proportions of looking to the target or any other IAs. Also, no random intercepts were necessary, as all dwell time intercepts would be close to the length of the IP.

For the pupil size model, random slopes for trial, text category, item type and funniness were added for the subject random factor. No random intercept for subject was implemented as we already z-standardized the pupil sizes beforehand.

In the provided R-Scripts on our OSF repository, the complete model specifications of the implemented models can be found.

The combined fixation reports from all participants contained 52,544 fixations. These were used to generate combined area of interest (IA) and period of interest (IP) reports, with each trial consisting of the summed dwell times for each of the three IAs and each of the two IPs, yielding 3×2 observations per trial and a total of $26 \times 32 \times 3 \times 2 = 4{,}992$ observations. Eighty trials (9.6%) with incorrect picture mapping responses were excluded, leaving a total of 4,512 (4,992–80 × 6) observations. This data set was used for the analysis of dwell time. For the pupil dilation data, 1803 observations were excluded if the dwell time for an area of interest was zero, leaving a total of 2,709 observations for pupil data analysis.

3 RESULTS

3.1 Behavioral Data

Estimated means and standard errors for the behavioural measures are shown in **Table 2**.

3.1.1 Funniness Ratings

The average funniness rating across all trials was 2.95 (SD = 2.36). The variance explained by the fixed effects in the model was 33.4%, including random effects raised this to 63.0%. The mixed-effects regression model showed a significant main effect for text category [$\chi2$ (1) = 92.80, $p < 0.0001$] and item type [$\chi2$ (1) = 177.44, $p < 0.0001$] as well as a significant interaction of these two

TABLE 2 | Estimated means and standard errors for the funniness ratings and the picture-matching accuracy for the four text categories.

	Joke		Revision	
	Experimental	Control	Experimental	Control
	M (SE)	M (SE)	M (SE)	M (SE)
Funniness rating (Scale 1-9)	5.22 (0.25)	2.19 (0.25)	2.76 (0.29)	1.64 (0.29)
Accuracy (%)	0.82 (0.06)	0.99 (0.01)	0.99 (0.001)	0.99 (0.01)

TABLE 3 | Statistical results of the analyses of the dwell times and the pupil sizes. Details of the models are described in the Results section.

Effect	df	Dwell times (p)	Pupil size (p)
text category	1	0.967	0.483
item type	1	0.503	0.060^
IP	1	0.015*	<0.001***
IA	2	0.001***	0.200
funniness	1	0.464	0.005**
trial	1	0.988	<0.001***
performance	1	0.202	0.451
Sense of humor	1	0.317	0.101
text category: item type	1	0.957	0.133
text category: IP	1	0.632	<0.001***
item type: IP	1	0.770	<0.001***
text category: IA	2	0.021*	0.003**
item type: IA	2	0.001***	0.834
IP: IA	2	<0.001***	0.519
text category: item type: IP	1	0.766	0.994
text category: item type: IA	2	0.665	0.273
text category: IP: IA	2	0.062^	0.624
item type: IP: IA	2	<0.001***	0.525
text category: item type: IP: IA	2	0.817	0.063^

*Text category = Joke/Revision; Item type = Experimental/Control; IA, interest area, Target/Competitor/Distractor; IP, interest period, before/after. Significance values: *** 0.001 ** 0.01 * 0.05 ^ 0.10.*

TABLE 4 | Statistical results of the analyses of the effects of funniness on the pupil sizes for the experimental texts. Details of the model are described in the Results section.

Effect	Chisq	Df	p
text category	0.4851	1	0.606
IP	76.0140	1	<0.001***
IA	1.3996	2	0.601
funniness	5.8860	1	0.055^
trial	36.9286	1	<0.001***
performance	0.0000	1	0.814
Sense of humour	4.8333	1	0.043*
text category: IP	0.0616	1	0.850
text category: IA	0.8284	2	0.663
IP: IA	2.7547	2	0.289
text category: funniness	0.0105	1	0.738
IP: funniness	10.9686	1	<0.001***
IA: funniness	3.7296	2	0.167
text category: IP: IA	2.8273	2	0.172
text category: IP: funniness	0.2846	1	0.572
text category: IA: funniness	0.3368	2	0.878
IP: IA: funniness	0.2320	2	0.778
text category: IP: IA: funniness	0.9381	2	0.671

*Text category = Joke/Revision; IA, interest area, Target/Competitor/Distractor; IP, interest period, before/after. Significance values: *** 0.001 ** 0.01 * 0.05 ^ 0.10.*

factors [$\chi2$ (1) = 37.77, p < 0.0001]. The model estimates in **Table 2** show that jokes yielded substantially higher funniness ratings than the other three text categories.

3.1.2 Picture Matching Task

In the picture matching task participants picked the correct target picture in 91.4% of the trials. In 7.3% of the trials the competitor picture was chosen, while the distractor was picked in only 2.3% of the trials. The logistic regression model explained 22.4% of the variance with fixed effects, and 61.5% with both random and fixed effects. The statistical analysis of the accuracy rates also showed a significant main effect for text category [$\chi2$ (1) = 11.85, p = 0.0006] and item type [$\chi2$ (1) = 14.86, p = 0.0001] as well as a significant interaction between the two factors [$\chi2$ (1) = 8.50, p = 0.0035]. The model estimates in **Table 2** indicate that while revision stories, revision control stories and joke control stories yielded a very high performance (99% estimated accuracy), the joke condition led to a higher error rate with an estimated accuracy of 82%.

3.2 Eye Movement Data

Results of the mixed-effects regression models for the dwell times and the average pupil size are presented in **Table 3**.

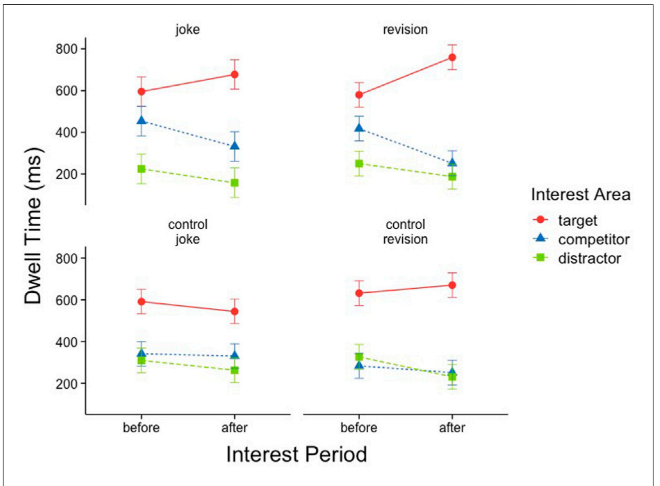

FIGURE 3 | Estimated mean dwell times for the four story types on target, competitor and distractor before and after the revision point. Error bars indicate the confidence interval.

3.2.1 Dwell Times

In the mixed-effects regression model for the analysis of the dwell times, the fixed effects explained 14.5% of the variance in dwell

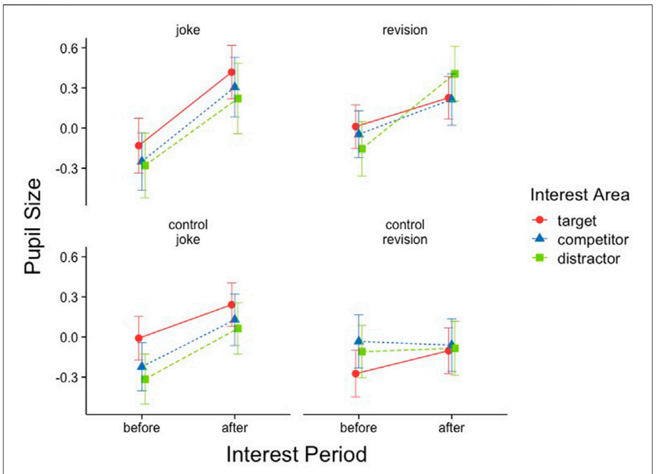

FIGURE 4 | Estimates of the grand-mean centred pupil size means for the four story types on target, competitor and distractor before and after the revision point.

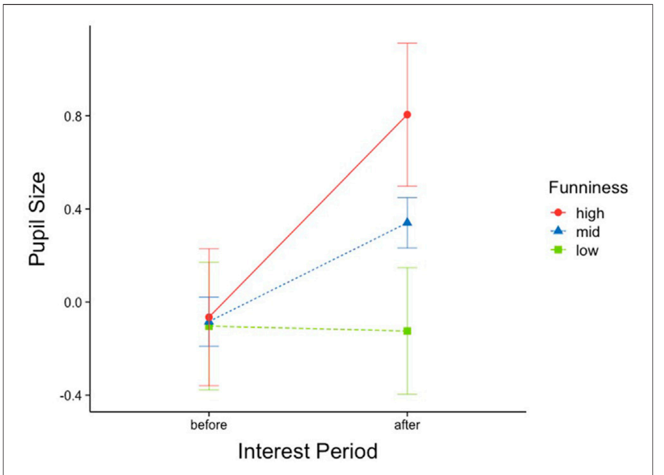

FIGURE 5 | Estimates of the grand—mean centred pupil size means as a function of interest period and the trial-based individual funniness ratings.

times while the model including random effects explained 24.2%. The Anova analysis (**Table 3**) showed significant main effects of IA and IP, as well as a significant two-way interaction of IA and IP. Inspection of the model estimates displayed in **Figure 3** shows that the target dwell times were higher than those of the distractor and competitor. This effect was more pronounced after the revision point than before. Crucially, the 2-way interaction of IA and item type, as well as a significant three-way interaction of IP, IA and item type indicate that the viewing patterns change mostly for the experimental texts (jokes and revision stories), but less so for the control texts.

Pairwise comparisons confirm this observation: Target dwell times increased for the experimental stories ($z = -4.3$, $p < 0.0001$), but not for the control stories ($z = 0.15$, n. s.); and competitor dwell times decreased for the experimental stories ($z = 4.7$, $p < 0.0001$), but not for the control stories ($z = 0.8$, n. s.). The distractor dwell times decreased for both item types (experimental stories: $z = 2.1$, $p < 0.05$; control stories: $z = 2.5$, $p < 0.05$).

The model also yielded a significant two-way interaction of text category and IA. Across both interest periods and averaged across experimental and control stories, there was a tendency for target pictures to be looked at longer for revision compared to joke stories ($z = -1.9$, $p = 0.06$), while the competitor pictures were viewed longer for the joke stories compared to the revision stories ($z = 2.4$, $p < 0.05$). There were no differences for the distractor pictures ($z = -0.38$, n. s.).

Finally, the results for the dwell times were independent of the trial position, the performance in the picture matching task, and the funniness ratings. None of the respective effects reached significance.

3.2.2 Pupil Size

The fixed effects of the multi-level regression model were able to explain 19.7% of the pupil size variance, fixed and random effects explained 37.9%. The Anova analysis over the model coefficients yielded a marginally significant main effect of IA, and a two-way interaction of IA with text category. The model estimates in **Figure 4** indicate that for the joke stories (both experimental and control), the pupil sizes were larger on the target pictures than on the competitor ($t = 2.4$, $p = 0.046$) and the distractor ($t = 3.2$, $p = 0.005$), with no difference between competitor and distractor ($t < 1$, n. s.). There was no effect of IA on the pupil sizes for the experimental and control revision stories (t's < 1, n. s.).

Importantly, the pupil sizes changed from before to after the revision point. The significant main effect for IP and an interaction of IP and text category was found. Inspection of the model estimates presented in **Figure 4** shows that experimental jokes and joke control texts elicited a larger increase of the pupil sizes than the revision and revision control stories taken together. Similarly, the main effect of item type and the interaction between item type and IP indicates larger pupil size increases for the experimental items than for the control texts. Despite the fact that both of these effects are driven by the lack of an increase in the revision control stories ($t < 1$, n. s.),—compared to highly significant increases for the other three conditions (|t|'s > 4.7, p's < 0.0001)—, the corresponding 3-way interaction between IP, text category and item type was far from significant.

The performance in the picture matching task did not influence pupil size. However, a main effect for trial (serial order) was present as mean pupil sizes decreased significantly over the course of the experiment. Most importantly, the perceived funniness of the trials influenced pupil dilation, as indicated by a significant main effect. In trials with higher funniness ratings, the participant's pupils were larger.

To investigate whether this effect was due to a reaction to the crucial information in the punchline, we conducted a further analysis for the experimental trials only (jokes and revision texts). The experimental texts elicited a wide range of funniness ratings, while over 85% of the control texts were rated very low (1-3). The model corresponds to the previous one but replaces the categorical variable item type by the funniness ratings. Its fixed effects account for 18.8% of the variance, including random effects yields 36.0%.

The results of this analysis are presented in **Table 4**. The effect of sense of humour indicates that participants who gave higher ratings overall had larger pupil sizes. The main effects of IP, funniness and trial were replicated. Furthermore, the highly significant interaction between funniness and IP confirmed that pupil sizes dilated more as a reaction to crucial information when the story was later rated as funnier. This effect is illustrated in **Figure 5**, where the continuous variable funniness is combined to three categories for sake of readability.

Text category did not have an effect, nor did interact with other variables. Thus, the distinction between experimental jokes and revision stories was not as important as the individually experienced funniness of the story.

4 DISCUSSION

In the present study a visual world paradigm was used to analyse the time course of situation model updating during the comprehension of verbal humour. Jokes were contrasted with non-humorous revision texts that also contained an incongruity, and with straightforward stories without the need for revision. The viewing patterns were comparable for jokes and the structurally identical revision stories.

4.1 Funniness Ratings and Picture Matching Performance

As in the studies of Siebörger (2006) and Ferstl et al. (2017) jokes were rated funnier than revision and control stories, despite the structural similarity of jokes and revision stories. This finding confirms that incongruency might be necessary for making a text funny, but that it is not sufficient. Thus, a differentiation of the cognitive and the affective aspects of humour processing became possible. The differences in funniness between the two text categories are consistent with differences in script opposition (Suls, 1972; Raskin, 1979). For example, in the revision story presented in **Table 1** the two narrative schemes only differ with respect to the agent of the scene (the cook or his wife preparing the meal). In the joke story on the other hand, Mark is introduced as a womanizer, but eventually turns out to be a thief. This more pronounced script opposition elicits an affective reaction.

Picture matching was expected to be a straightforward task. And indeed, it did not pose any problems for the revision stories and the coherent control stories. Surprisingly, however, the accuracy was considerably lower in the joke condition. In previous reading studies with similar materials, the revision stories had been most difficult (e.g., Ferstl et al., 2017). One explanation for this finding is that in some jokes the alternative situations reflect different perspectives on the same state of affairs, i.e., the joke is based on a misunderstanding between two protagonists. Consider the following example:

> The mother explains to her son: Today mom and dad are married for 10 years. The son: And for how much longer do you have to stay together?

In this joke, the correct choice in the picture matching task is the meaning implicated by the punchline: the son views his parents' marriage as an unpleasant commitment. However, the funniness of the story depends on the incompatibility of the son's attitude with the attitude of his mother who regards her anniversary as a joyful event. Thus, the initial situation remains active during the processing of the punchline, so that both the competitor and the target are plausible choices in the picture matching task.

To control for the individual differences in picture matching performance incorrect trials were excluded and the participant's picture matching performance for jokes was added as a continuous covariate in all statistical analyses. As this covariate did not yield any significant effects, we are confident that the eye tracking results were independent of task induced strategies.

4.2 Eye Tracking: Viewing Times

The dwell times for the unambiguous coherent control texts showed that, as expected, the differentiation between the target picture and the competitor was completed before the revision point in the punchline. Because only the target picture was relevant for the interpretation of these stories, the competitor picture was viewed as little as the distractor picture. As the punchline did not contain any incongruous information, there were no systematic changes in the fixations patterns at the revision point. In the jokes and revision stories on the other hand, participants fixated both the competitor and the target pictures before the crucial information was presented at the revision point. A rise of target fixations and a switch from the initial to the final meaning of the story occurred within 1,500 ms after the revision point. Using ERPs, Canal et al. (2019) found that the processing sequence consisting of incongruency detection, incongruency resolution and later interpretation was completed within this timeframe.

The time window is also in line with a study by Fiacconi and Owen (2015) who located the moment of insight, i.e., the instant a verbal joke is understood, at around 800 ms after reading the punchline. In their study participants also heard jokes and unfunny but ambiguous control texts with an incongruity arising at the very end of each story. Examples are:

> Joke: What did the teddy bear say when he was offered a dessert? No thank you, I'm stuffed!
> Control: What was the problem with the other coat? It was very difficult to put it on with the paint roller!

EMG measurements on the musculus zygomaticus showed that first indications of a smile, interpreted an evidence for joke appreciation, were found 800 ms after reading the punchline of jokes. Fiacconi and Owen (2015) also reported a cardiovascular reaction about 5,000 ms after the punchline that correlated with the perceived funniness.

It would be desirable to describe the temporal sequence of processing in a more fine-grained way. However, the exact timecourse depends on properties of the texts and the experimental procedure. Independent of the funniness of the texts, situation model building and inferencing, as required for the comprehension of texts, depend on the exact wording, the context and the experimental setting.

The fact that target fixations in experimental jokes and revision stories in the present study were already high before the revision point indicates that the correct meaning was anticipated before the occurrence of the incongruity. This anticipation was possibly elicited by the visual material that depicted both alternative interpretations.

The viewing patterns of the two types of experimental texts were indistinguishable. There was no evidence for faster incongruity resolution or longer target dwell times in jokes. Although this null effect cannot be interpreted as providing evidence for similarity—but only as a lack of difference in this particular experiment—it is consistent with the interpretation that the linguistic revision drives the viewing pattern. This result, together with the finding of comparable first pass reading times in the previous study, suggests that the observed differences between these texts in self-paced reading might have reflected meta-cognitive strategies. The affective reaction in jokes presumably provided intrinsic feedback about the correctness of the incongruity resolution, leading to shorter overall reading times for jokes compared to revision stories in the previous reading study (Ferstl et al., 2017).

4.3 Eye Tracking: Pupil Dilation

The analysis of the pupil sizes added important information to the viewing patterns. Higher mean pupil sizes were present on the target picture compared to competitor and distractor pictures for experimental jokes and revision stories. Somewhat surprisingly, a larger pupil size on the target IA was also observed for joke control texts that neither contained an incongruity nor a funny punchline that could account for an increase in pupil size on the target.

The revision control texts, on the other hand, showed pupil size means below the participant's average for all three IA. As all control texts were created in exactly the same way, by changing the context of the experimental text to activate the final situation model, these differences are likely to be due to the visual displays. Although the visual salience was comparable, the content of the target and competitor scenes created for the joke texts reflected the script opposition and were thus more distinct, and possibly more interesting, than those created for the revision texts. It is also possible that the picture material used for the experimental jokes as well as for the matching joke control texts was more funny in general, triggering an affective reaction for the control stories even though the corresponding text did not contain a funny punchline.

The two-way interactions of interest period with text category and item type, respectively, indicated that pupil sizes increased to some extent for all texts, except for the revision control texts. This finding is in line with a sensitivity to revision demands, as well as to affective aspects. And in fact, the main effect of the funniness ratings on the pupil sizes during text comprehension confirms that changes in pupil size were influenced by content that the participants reported to elicit affective reactions (e.g., Partala and Surakka, 2013). The pupils dilated more for funnier stories. An additional analysis on the experimental trials confirmed that the pupil dilation varied with the funniness of the story, rather than with the predefined text category. A similar result was reported by Mayerhofer and Schacht (2015) who measured pupil diameters during the reading of jokes and control texts. The pupil sizes increased about 800 ms after humourous endings, compared to coherent control texts. Similar to our findings, this effect correlated with the funniness of the jokes, as determined in independent ratings.

The estimated mean pupil sizes decreased over the course of the experiment which could indicate that the cognitive load decreased as participants got used to the task requirements and stories. Moreover, as participants might have anticipated the punchline of jokes more strongly over the course of the experiment based on these learning effects, the strength of the affective reaction might have decreased as well, as the funniness response in jokes is assumed to be a result of the surprising occurrence of the incongruity (Dynel, 2009; Canal et al., 2019).

4.4 The Visual World Paradigm

The present study confirms that the visual world paradigm is a useful tool for studying situation model updating, and, in particular, the processing of humorous language. The fine-grained temporal resolution of the viewing patterns allowed us to describe the time course of comprehension in detail. Auditory presentation is a rather natural presentation modality, and reading strategies did not play a role. Finally, appropriate control conditions eliminated lexical or sentence level effects and deconfounded joke appreciation and cognitive revision. Moreover, the analysis of pupil sizes provided additional information about the affective component of humour processing. The perceived funniness of the story was accompanied by larger pupil dilations, and joke and joke control texts—sharing more interesting picture displays—also elicited larger increases in pupil sizes. Applying this method to other issues in humor comprehension research can aid in further developing theories of humor processing that take into account the interplay between cognitive demands and affective reactions.

4.5 Limitations

Although the visual world paradigm is clearly appropriate for studying verbal humour, the unexpected differences between the two types of control texts indicate that the content of the visual displays influenced processing of the verbal input which was presented later. Further research is needed to understand the interplay between features of the visual scenes with the auditory language input during higher level language comprehension (Hüttig et al., 2011).

A further limitation of the present study is the small number of items used. An increase in experimental power would be desirable. However, the number of eight trials per condition is not unusual for visual world experiments. In addition to the constraints for designing materials for studies on joke comprehension, a visual world study requires visualizable texts for which both competing interpretations can be visualized in comparable pictures.

Nevertheless, due to the repeated-measures design of the study, a substantial number of observations was collected for the analysis of the behavioural and eye-tracking data. Because of the rather small number of observations on the subject level though, we could not explore the effects of gender on the eye-tracking dynamics of joke comprehension. As effects of gender on joke comprehension have been reported in previous studies though (Ferstl et al., 2017), future studies should also investigate these.

Because of the complex experimental design, the hierarchical structure of the repeated-measures data and the requirements of the visual world paradigm, we implemented a simplified analysis of the temporal sequence of fixations. This choice allowed us to pinpoint the effects of the information presented at the revision point. A more fine-grained analysis of the timecourse (for example using growth curve analysis, Mirman et al., 2008) would enable us to compare the present results to those using different methods, such as ERPs (Canal et al., 2019) or EMG (Fiacconi and Owen, 2015). The present study, however, can be viewed as a first exploratory analysis of the temporal progression of joke comprehension using the visual world paradigm and can be used for generating new hypotheses in this new field of research.

5 CONCLUSION

The present visual world study showed that jokes and non-funny revision texts involve similar cognitive processes. The viewing patterns confirmed that after the occurrence of an incongruity the initial situation model is revised und replaced by the globally correct one. As there was no facilitation effect for jokes, the processing advantages for jokes found in reading studies are likely to be due to meta-cognitive evaluation processes. Importantly, pupil size analyses can shed light on the affective component of the processing of verbal humour. The combination of these information sources from visual world experiments provides a promising tool for studying the interplay between cognitive and affective aspects of humor processing.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in the OSF repository (https://osf.io/qk95j/).

AUTHOR CONTRIBUTIONS

The experiment was conducted by LI in partial fulfilment of the requirements of a M. Sc. degree in Cognitive Science at the Albert-Ludwigs-University in Freiburg. All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

REFERENCES

Adobe Inc. (2019). *Adobe Illustrator*. Retrieved from https://adobe.com/products/illustrator.

Altmann, G. T. M., and Kamide, Y. (2007). The Real-Time Mediation of Visual Attention by Language and World Knowledge: Linking Anticipatory (And Other) Eye Movements to Linguistic Processing. *J. Mem. Lang.* 57, 502–518. doi:10.1016/j.jml.2006.12.004

Altmann, G. T. M. (2011). “The Mediation of Eye Movements by Spoken Language,” in *The Oxford Handbook of Eye Movements*. Editors S. P. Liversedge, I. D. Gilchrist, and S. Everling. Oxford University Press, 979–1003.

Attardo, S., and Raskin, V. (1991). Script Theory Revis(it)ed: Joke Similarity and Joke Representation Model. *Humour* 4 (3), 293–347. doi:10.1515/humr.1991.4.3-4.293

Attardo, S. (1997). The Semantic Foundations of Cognitive Theories of Humor. *Humour* 10 (4), 395–420. doi:10.1515/humr.1997.10.4.395

Barr, D. J. (2008). Analyzing 'visual World' Eyetracking Data Using Multilevel Logistic Regression. *J. Mem. Lang.* 59, 457–474. doi:10.1016/j.jml.2007.09.002

Bates, D., Maechler, M., Bolker, B., and Walker, S. (2014). lme4: Linear Mixed-Effects Models Using Eigen and S4. R Package Version 1.1-7. Available at: http://CRAN.R-project.org/package=lme4.

Berends, S. M., Brouwer, S. M., and Sprenger, S. A. (2015). “Eye-tracking and the Visual World Paradigm,” in *Designing Research on Bilingual Development: Behavioural and Neurolinguistic Experiments*. Editor M. Schmid, et al. (Cham: Springer), 55–80. doi:10.1007/978-3-319-11529-0_5

Bradley, M. M., Codispoti, M., Cuthbert, B. N., and Lang, P. J. (2001). Emotion and Motivation I: Defensive and Appetitive Reactions in Picture Processing. *Emotion* 1, 276–298. doi:10.1037/1528-3542.1.3.276

Bradley, M. M., Miccoli, L., Escrig, M. A., and Lang, P. J. (2008). The Pupil as a Measure of Emotional Arousal and Autonomic Activation. *Psychophysiology* 45, 602–607. doi:10.1111/j.1469-8986.2008.00654.x

Bressler, E. R., and Balshine, S. (2006). The Influence of Humor on Desirability. *Evol. Hum. Behav.* 27 (1), 29–39. doi:10.1016/j.evolhumbehav.2005.06.002

Buchowski, M. S., Majchrzak, K. M., Blomquist, K., Chen, K. Y., Byrne, D. W., and Bachorowski, J. A. (2007). Energy Expenditure of Genuine Laughter. *Int. J. Obes. (Lond)* 31, 131–137. doi:10.1038/sj.ijo.0803353

Canal, P., Bischetti, L., Di Paola, S., Bertini, C., Ricci, I., and Bambini, V. (2019). 'Honey, Shall I Change the Baby? - Well Done, Choose Another One': ERP and Time-Frequency Correlates of Humor Processing. *Brain Cogn.* 132, 41–55. doi:10.1016/j.bandc.2019.02.001

Cooper, R. M. (1974). The Control of Eye Fixation by the Meaning of Spoken Language. *Cogn. Psychol.* 6, 84–107. doi:10.1016/0010-0285(74)90005-x

Coulson, S. (2006). Constructing Meaning. *Metaphor and Symbol* 21 (4), 245–266. doi:10.1207/s15327868ms2104_3

Coulson, S., and Kutas, M. (1998). Frame-shifting and Sentential Integration. *UCSD Cogn. Sci. Tech. Rep.* 98 (3), 1–32.

Coulson, S., and Kutas, M. (2001). Getting it: Human Event-Related Brain Response to Jokes in Good and Poor Comprehenders. *Neurosci. Lett.* 316, 71–74. doi:10.1016/s0304-3940(01)02387-4

Coulson, S., and Lovett, C. (2004). Handedness, Hemispheric Asymmetries, and Joke Comprehension. *Cogn. Brain Res.* 19, 275–288. doi:10.1016/j.cogbrainres.2003.11.015

Coulson, S., Urbach, T., and Kutas, M. (2006). Looking Back: Joke Comprehension and the Space Structuring Model. *Humour* 19 (3), 229–250. doi:10.1515/humor.2006.013

Coulson, S., and Williams, R. F. (2005). Hemispheric Asymmetries and Joke Comprehension. *Neuropsychologia* 43, 128–141. doi:10.1016/j.neuropsychologia.2004.03.015

Dynel, M. (2009). Beyond a Joke: Types of Conversational Humour. *Lang. Linguistics Compass* 3 (9), 1284–1299. doi:10.1111/j.1749-818x.2009.00152.x

Engelhardt, P. E., Ferreira, F., and Patsenko, E. G. (2010). Pupillometry Reveals Processing Load during Spoken Language Comprehension. *Q. J. Exp. Psychol.* 63, 639–645. doi:10.1080/17470210903469864

Experiment Builder [Computer software] (2014). Available at: http://www.sr-research.com.

EyeLink Data Viewer User's Manual (2011). *Document Version 1.11.1*. Mississauga, ON, Canada: SR Research Ltd.

Ferstl, E. C., Israel, L., and Putzar, L. (2017). Humor Facilitates Text Comprehension: Evidence from Eye Movements. *Discourse Process.* 54 (4), 259–284. doi:10.1080/0163853X.2015.1131583

Ferstl, E. C., and Kintsch, W. (1999). "Learning from Text: Structural Knowledge Assessment in the Study of Discourse Comprehension," in *The Construction of Mental Models during reading*. Editors H. Oostendorp, and S. Goldman (Mahwah, NJ: Lawrence Erlbaum), 247–277.

Fiacconi, C. M., and Owen, A. M. (2015). Using Psychophysiological Measures to Examine the Temporal Profile of Verbal Humor Elicitation. *PLoS ONE* 10 (9), e0135902. doi:10.1371/journal.pone.0135902

Fox, J., and Weisberg, S. (2011). *An {R} Companion to Applied Regression*. 2nd ed. Thousand Oaks CA: Sage. Available at: http://socserv.socsci.mcmaster.ca/jfox/Books/Companion.

Giora, R. (1991). On the Cognitive Aspects of the Joke. *J. Pragmatics* 16, 465–485. doi:10.1016/0378-2166(91)90137-m

Giora, R. (2002). "Optimal Innovation and Pleasure," in *Processing of the April Fools' Day Workshop on Computational Humour. Trento, Itlay*. Editors O. Stock, C. Strapparva, and A. Nijholt (Trento, Italy: ITC-itst), 20 11–28.

Goel, V., and Dolan, R. J. (2001). The Functional Anatomy of Humor: Segregating Cognitive and Affective Components. *Nat. Neurosci.* 4, 237–238. doi:10.1038/85076

Grice, H. (1975). "Logic and Conversation," in *Syntax and Semantics*. Editors P. Cole, and J. Morgan (New York: Academic Press), 3, 45–58.

Hess, E. H., and Polt, J. M. (1960). Pupil Size as Related to Interest Value of Visual Stimuli. *Science* 132, 349–350. doi:10.1126/science.132.3423.349

Huettig, F., and Altmann, G. T. (2005). Word Meaning and the Control of Eye Fixation: Semantic Competitor Effects and the Visual World Paradigm. *Cognition* 96, B23–B32. doi:10.1016/j.cognition.2004.10.003

Huettig, F., Olivers, C. N. L., and Hartsuiker, R. J. (2011). Looking, Language, and Memory: Bridging Research from the Visual World and Visual Search Paradigms. *Acta Psychologica* 137, 138–150. doi:10.1016/j.actpsy.2010.07.013

Hunger, B., Siebörger, F., and Ferstl, E. C. (2009). Schluss mit lustig: Wie Hirnläsionen das Humorverständnis beeinträchtigen [Done with fun: How brain lesions affect humour comprehension]. *Neurolinguistik* 21, 35–59.

Hüttig, F., Rommers, J., and Meyer, A. S. (2011). Using the Visual World Paradigm to Study Language Processing: A Review and Critical Evaluation. *Acta Psychologica* 137, 151–171. doi:10.1016/j.actpsy.2010.11.003

Hyönä, J., Tommola, J., and Alaja, A. (2007). Pupil Dilation as a Measure of Processing Load in Simultaneous Interpretation and Other Language Tasks. *Q. J. Exp. Psychol.* 48 (A), 598–612. doi:10.1080/14640749508401407

Kant, I. (1790/1977). Kritik der Urteilskraft. *Werkausgabe Band X*. Frankfurt am Main: Suhrkamp Verlag.

Kimata, H. (2004). Laughter Counteracts Enhancement of Plasma Neurotrophin Levels and Allergic Skin Wheal Responses by mobile Phone-Mediated Stress. *Behav. Med.* 29, 149–154. doi:10.3200/bmed.29.4.149-154

Lenth, R. (2014). Lsmeans: Least-Squares Means. R Package Version 2.10. Available at: http://CRAN.R-project.org/package=lsmeans.

Martin, A. (2007). *The Psychology of Humour: An Integrative Approach*. Burlington: Elsevier Academic Press.

Mayerhofer, B. (2015). *Universität Göttingen*. Unpublished dissertation. Available at: http://hdl.handle.net/11858/00-1735-0000-0022-5DC9-4.Processing of Garden Path Jokes: Theoretical Concepts and Empirical Correlates

Mayerhofer, B., and Schacht, A. (2015). From Incoherence to Mirth: Neuro-Cognitive Processing of Garden-Path Jokes. *Front. Psychol.* 6, 550. doi:10.3389/fpsyg.2015.00550

Mirman, D., Dixon, J. A., and Magnuson, J. S. (2008). Statistical and Computational Models of the Visual World Paradigm: Growth Curves and Individual Differences. *J. Mem. Lang.* 59 (4), 475–494. doi:10.1016/j.jml.2007.11.006

Mitchell, H. H., Graesser, A. C., and Louwerse, M. M. (2010). The Effect of Context on Humor: A Constraint-Based Model of Comprehending Verbal Jokes. *Discourse Process.* 47, 104–129. doi:10.1080/01638530902959893

Partala, T., and Surakka, V. (2003). Pupil Size Variation as an Indication of Affective Processing. *Int. J. Human-Computer Stud.* 59, 185–198. doi:10.1016/s1071-5819(03)00017-x

Pexman, P. M., Rostad, K. R., McMorris, C. A., Climie, E. A., Stowkowy, J., and Glenwright, M. R. (2011). Processing of Ironic Language in Children with High-Functioning Autism Spectrum Disorder. *J. Autism Dev. Disord.* 41 (8), 1097–1112. doi:10.1007/s10803-010-1131-7

Prehn, K., Kazzer, P., Lischke, A., Herpertz, S. C., Lischke, A., Heinrichs, M., et al. (2013). Effects of Intranasal Oxytocin on Pupil Dilation Indicate Increased Salience of Socioaffective Stimuli. *Psychophysiol* 50, 528–537. doi:10.1111/psyp.12042

Pyykkönen, P., Hyönä, J., and van Gompel, R. P. G. (2010). Activating Gender Stereotypes during Online Spoken Language Processing: Evidence from Visual World Eye-Tracking. *Exp. Psychol.* 57, 126–133.

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at: https://www.R-project.org/.

Raskin, V. (1979). Semantic Mechanisms of Humor. *Bls* 5, 325–335. doi:10.3765/bls.v5i0.2164

Salverda, A. P., and Tanenhaus, M. K. (2018). "The Visual World Paradigm," in *Research Methods in Psycholinguistics and the Neurobiology of Language: A Practical Guide*. Editors A. M. B. de Groot, and P. Hagoort (Malden, MA: Wiley-Blackwell), 89–110.

Schmidt, S. R. (2002). The Humour Effect: Differential Processing and Privileged Retrieval. *Memory* 10, 127–138. doi:10.1080/09658210143000263

Searle, J. (1979). *Expression and Meaning: Studies in the Theory of Speech Acts*. Cambridge: Cambridge University Press.

Siebörger, F. T. (2006). *Funktionelle Neuroanatomie des Textverstehens. Kohärenzbildung bei Witzen und anderen ungewöhnlichen Texten [*Functional neuroanatomy of text comprehension: Coherence building in jokes and other unusual texts*]*. Dissertation. Universität Leipzig, Leipzig: MPI Series in Human Cognitive and Brain Sciences.

Sirois, S., and Brisson, J. (2014). Pupillometry. *Wires Cogn. Sci.* 5, 679–692. doi:10.1002/wcs.1323

Strick, M., Holland, R. W., van Baaren, R., and van Knippenberg, A. (2010). Humor in the Eye Tracker: Attention Capture and Distraction from Context Cues. *J. Gen. Psychol.* 137, 37–48. doi:10.1080/00221300903293055

Suls, J. M. (1972). "A Two-Stage Model for the Appreciation of Jokes and Cartoons: An Information-Processing Analysis," in *The Psychology of Humour. Theoretical Perspectives and Empirical Issues*. Editors P. Goldstein, and J. McGhee (New York: Academic Press), 81–100. doi:10.1016/b978-0-12-288950-9.50010-9

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., and Sedivy, J. C. (1995). Integration of Visual and Linguistic Information in Spoken Language Comprehension. *Science* 268 (5217), 1632–1634. doi:10.1126/science.7777863

van Dijk, T. A., and Kintsch, W. (1983). *Strategies of Discourse Comprehension*. New York, NY: Academic Press.

Volkmann, B., Siebörger, F., and Ferstl, E. C. (2008). *Spass Beiseite? [All Joking Aside?* Hofheim: NAT-Verlag.

Winn, B., Whitaker, D., Elliott, D. B., and Phillips, N. J. (1994). Factors Affecting Light-Adapted Pupil Size in Normal Human Subjects. *Invest. Ophthalmol. Vis. Sci.* 35 (3), 1132–1137.

Wong, A. Y., Moss, J., and Schunn, C. D. (2016). Tracking reading Strategy Utilisation through Pupillometry. *Ajet* 32 (6), 45–57. doi:10.14742/ajet.3096

Zwaan, R. A., Stanfield, R. A., and Yaxley, R. H. (2002). Language Comprehenders Mentally Represent the Shapes of Objects. *Psychol. Sci.* 13 (2), 168–171. doi:10.1111/1467-9280.00430

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.

ORIGINAL RESEARCH
published: 16 March 2022
doi: 10.3389/fpsyg.2022.823313

Linguistic Features and Psychological States: The Case of Virginia Woolf

*Xiaowei Du**

Department of Foreign Language, Huazhong University of Science and Technology, Wuhan, China

This study investigated the relation between psychological states and linguistic features with the case of Virginia Woolf. We analyzed the data from *The Diary of Virginia Woolf* and *Virginia Woolf: Biography* by automatic text analysis and statistical analysis, including stepwise multiple regression and Deep Learning algorithm. The results suggested that the significant linguistic features can jointly predict the psychological states of Virginia Woolf, including the emotional value of *anger*, the absolutist word "*everything*," and the total of first-person plural pronouns. In addition, we found that the total use of first-person plural pronouns and the emotional value of *anger* were negatively related to mental health of Virginia Woolf. While the use of the absolutist word "*everything*" was positively related to mental health of Virginia Woolf. Meanwhile, we developed a model that can predict the psychological states of Virginia Woolf, with 86.9% accuracy. We discussed the findings and enumerated the limitations of this study at the end of the paper. The results not only complemented previous studies in the understanding of the relation between language and psychological health, but also facilitated timely identification, intervention, and prevention of mental disorders.

Keywords: psychological states, linguistic features, prediction, mental disorders, prevention

OPEN ACCESS

Edited by:
*Michael J. Cortese,
University of Nebraska Omaha,
United States*

Reviewed by:
*Concetta Papapicco,
University of Bari Aldo Moro, Italy
Franca Crippa,
University of Milano-Bicocca, Italy*

Correspondence:
*Xiaowei Du
d201981214@hust.edu.cn
orcid.org/0000-0001-8354-246X*

Specialty section:
*This article was submitted to
Language Sciences,
a section of the journal
Frontiers in Psychology*

Received: *27 November 2021*
Accepted: *26 January 2022*
Published: *16 March 2022*

Citation:
Du X (2022) Linguistic Features and Psychological States: The Case of Virginia Woolf. Front. Psychol. 13:823313. doi: 10.3389/fpsyg.2022.823313

INTRODUCTION

The relation between psychological states and linguistic features recently has attracted the interest of researchers. The motivation of such a line of research is that the psychological states of a person can be examined by his/her linguistic features (Nguyen et al., 2017) since the language of a person seems a reliable device to interpret his/her internal thoughts, emotions, and feelings (Herbert et al., 2019).

Some studies have explored the relation between psychological states and linguistic features. These studies are roughly categorized into two types. First, psychological states are assessed by the way people write about their experiences (Barnes et al., 2007; Tausczik and Pennebaker, 2010; Al-Mosaiwi and Johnstone, 2018; Kim et al., 2019). Second, psychological states are predicted by factors such as the diachronic changes of linguistic features in written texts (Rodrigues et al., 2016; Ziemer and Korkmaz, 2017; Eichstaedt et al., 2018; Boukil et al., 2019).

However, mixed findings were obtained regarding the relation between linguistic features and psychological states. First, psychological states are found developing closely related with different linguistic feature. For example, Barnes et al. (2007) found that the suicide texts reflect a linguistic trend in negative emotion words and death words. In contrast, Kim et al. (2019) showed a linguistic trend in modifier, numerals, first and second person pronouns, emotion words (positive, negative, sadness, and depression-related), and future tense verbs in the suicide texts. Second, the relation between linguistic features found in texts and psychological states are inconsistent. For example, Tausczik and Pennebaker (2010) found that people who consider suicide express more negative emotional words. Another example showed an opposite result that people who consider suicide express less negative emotional words (Barnes et al., 2007).

Therefore, the purpose of this study is to explore the relation between psychological states and linguistic features with the case of Virginia Woolf. The findings of this study may examine and complement previous studies and provide an approach to early identification, intervention, and prevention of mental disorders.

LITERATURE REVIEW

In this section, we introduce linguistic features that have been used to recognize psychological states in previous studies and a computational method that has been used to analyze sentiments and emotions.

Linguistic Features and Psychological States

Recent studies have shown that linguistic features such as personal pronouns, emotion words, absolutist words, color words, word count, and question marks may be used to identify psychological states.

The use of personal pronouns has served as a promising indicator of psychological states (Tausczik and Pennebaker, 2010). For example, Campbell and Pennebaker (2003) suggested that personal pronouns manifest psychological states from the perspective of a person's social connection or social isolation. Specifically, an excessive use of first-person singular pronouns may be related to a high degree of self-involvement, while an increased use of the other pronouns may indicate improvement of social engagement (Cohn et al., 2004; Simmons et al., 2008). In addition, people use more first-person singular pronouns when in grief or depression or attempting suicide (Rude et al., 2004; Boals and Klein, 2005; Eichstaedt et al., 2018). Last, it is also found that fewer first-person singular pronouns may involve deceptive communications (Newman et al., 2003), while more first-person plural pronouns may indicate that he or she was in a happy marriage (Simmons et al., 2008).

Similarly, the use of emotional words is also correlated with psychological states (Barnes et al., 2007), though mixed findings were found regarding the relation between emotions and psychological states. The findings can be roughly categorized into three types. First, both positive and negative emotional words are related with mental health, and more negative emotional words and less positive emotional words reflect a less healthy mental state (Pennebaker et al., 2003; Kahn et al., 2007). Second, only negative emotional words are related with mental health, and more negative emotional words reflect a less healthy mental health (Kahn et al., 2007; Herbert et al., 2019). Third, more emotional expressions improve psychological states, such as positive emotional words, anger, or sadness (Rude et al., 2004; Graves et al., 2005; Barnes et al., 2007). It is worth noting that some studies stressed the impact of negative emotional expression on psychological states. Previous studies found that negative emotional words not only carry more information about mental health than positive emotional words (Garcia et al., 2012) but may also benefit psychological states (Lerner et al., 2003).

Absolutist words have also been employed to predict psychological states (Savekar et al., 2019) since absolutist words are believed to present an absolutist thinking (Al-Mosaiwi and Johnstone, 2018). Empirical studies have revealed that the absolutist thinking, a cognitive distortion related to extreme and rigid thoughts, may do harm to mental health (Savekar et al., 2019; Jones et al., 2020). More importantly, compared with negative emotional words, absolutist words may more accurately track the severity of the affective disorder (Al-Mosaiwi and Johnstone, 2018). To be specific, the language of suicidal ideation contains approximately 30% more absolutist words than that of anxiety and depression, and approximately 80% more than that of normal mental health (Al-Mosaiwi and Johnstone, 2018).

Other linguistic features such as color words, question marks, and word counts, may also be used to predict psychological states (Barnes et al., 2007; Tausczik and Pennebaker, 2010; Wadsworth et al., 2016). For example, Wadsworth et al. (2016) found that the use of *white* in Sylvia Plath's poems increased significantly while the use of *gray*, *yellow*, *purple*, and *brown* decreased rapidly before Sylvia Plath committed suicide. Furthermore, Barnes et al. (2007) analyzed letters and diaries of suicidal youth and found that the use of question marks increased rapidly before their suicide. Last, some studies indicated that word counts may be related to psychological states since it reflects how engaged people are in the expression or how much information people produced (Alvarez-Conrad et al., 2001; Tausczik and Pennebaker, 2010; Wadsworth et al., 2016).

Sentiment Analysis and Psychological States

Sentiment analysis has recently been used to study psychological states since it could take a step closer to understanding and anticipating an individual's physiological states and needs (Al-Thubaity et al., 2018; Chatterjee et al., 2019; Moreno-Blanco et al., 2020). To be specific, sentiment analysis quantifies or extracts the polarity of sentiments, attitudes, emotions, and opinions of a given text (Cambria, 2016; Rendalkar and Chandankhede, 2018). More importantly, sentiment analysis broadly covers two dimension, namely, sentiment analysis in a narrow sense and emotion analysis (Lei and Liu, 2021). On the one hand, sentiment analysis in a narrow sense mainly identifies the sentiment polarities such as positive, negative, or neutral (D'Andrea et al., 2015; Lennox et al., 2020; Zucco et al., 2020). Some studies have applied sentiment analysis in a narrow sense to identify psychological states (Tausczik and Pennebaker, 2010). For example, Herbert et al. (2019) found that people express more negative emotions in their notes before committing suicide. Similarly, Eichstaedt et al. (2018) found that depressed people express more negative emotions, which

TABLE 1 | The classification of variables.

The dependent variable	Psychological states	*1 normal, 2 recovery, 3 restlessness, anxiety, and irritability 4 physical discomfort, 5 mental illness*
The independent variables	Colors	*red, white, black, green, yellow, blue, purple, gray*
	Personal pronouns	*i, my, me, we, us, our, she, her, he, him, his, it, its, them, their, they, you, your, yours, i_my_me, we_our_us, you_your_yours, he_his_him, she_her, it_its, they_them_their*
	Absolutist word	*absolutely, all, always, entire, ever, every, everyone, everything, definitely, full, must, complete, completely, constant, constantly, whole never, nothing, totally, total absolutist words*
	Emotions	*anticipation, sadness, joy, trust, disgust, anger, surprise, fear*
	Question mark	
	Word counts	
	Sentiment value	

contributed negatively to their psychological states. In addition, Rodrigues et al. (2016) detected positive or negative emotions of cancer patients by analyzing their texts in online communities to support their treatment.

On the other hand, emotion analysis focuses on recognizing a set of basic emotions, which may extend a more accurate and reliable detection of psychological states and feelings (Calabrese and Cannataro, 2016; Molina Beltrán et al., 2019). Previous studies have developed several versions of basic emotions. Specifically, Ekman (1993) proposed six universal emotions including anger, disgust, fear, sadness, joy, and surprise. Besides, Plutchik (1980) developed eight basic emotions such as anger, anticipation, disgust, fear, joy, sadness, surprise, and trust. Some studies have investigated the relation between basic emotions and psychological states (Desmet and Hoste, 2013; Ciullo et al., 2016). For instance, suicide individuals tend to express more emotional words such as anxiety, sadness, and depression in response to their feelings of anger or anxiety (Kim et al., 2019). Meanwhile, the use of joy or happiness words, revealing a sense of enjoyment, satisfaction, and pleasure, and these words are frequently used when an individual is in the situation of wellbeing, inner peace, love, safety, and contentment (Papapicco and Mininni, 2020). Additionally, the use of sadness words reflects the degree of social withdrawal or mood flattening, occurring with a higher frequency when an individual is most likely in grief, loss, frustration, depression, and suicide ideation (Barnes et al., 2007; Eichstaedt et al., 2018; Kim et al., 2019).

In fact, many studies have explored the relation between psychological states and linguistic features. However, a few limitations should be noted. First, most studies applied commercial tools such as Linguistic Inquiry and Word Count Program (LIWC) to analyze the psychological states of texts, which may limit the application of their findings in that such tools are not freely available. More importantly, these studies can only explore the linguistic features provided by LIWC, which may limit the comprehensiveness of the research. Second, few studies have examined the relation between psychological states and other linguistic features besides personal pronouns and positive and negative emotions (Wadsworth et al., 2016; Al-Mosaiwi and Johnstone, 2018). Third, previous studies yielded conflicting findings, such as different linguistic features that are related to psychological states (Eichstaedt et al., 2018; Kim et al., 2019), and a different relation between linguistic features and psychological states (Barnes et al., 2007; Tausczik and Pennebaker, 2010).

Therefore, this study aims to address the foregoing limitations and examine the relation between linguistic features and psychological states with the case of Virginia Woolf. More specifically, we use *Python* and *R* (version 3.6.0) to analyze the psychological states of texts, which can consider more linguistic features. In addition, this study examine linguistic features that are related to psychological states of Virginia Woolf. We also confirm the relation between linguistic features and the psychological states of Virginia Woolf. Finally, we aim to provide an equation to predict the psychological states of Virginia Woolf by linguistic features.

MATERIALS AND METHODS

In this section, we introduce the data, the methods for text analysis, and the statistical analysis in this study.

Data

The Diary of Virginia Woolf (Bell, 1984) was used for text analysis in the present study for three reasons. First, Virginia Woolf was a British novelist and essayist regarded as one of the major modernist literary figures of the twentieth century (Boeira et al., 2016). Second, she had mental disorders and attempted suicide several times, and killed herself at the age of 59 (Androutsopoulou et al., 2019). Third, *The Diary of Virginia Woolf* (Bell, 1984) contains 1,577 diaries in total and covers 26 years, from 1915 until 4 days before death in 1941, except in 1916, Virginia Woolf didn't write a diary due to a serious breakdown (Blodgett, 1989). It is a private text of self-presentation and provides more comprehensive and affluent information to reveal the relation between linguistic features and psychological states (Briggs, 2011; Androutsopoulou et al., 2019). We saved each diary piece as one text and obtained 1,577 texts in total.

Virginia Woolf: A Biography (Bell, 1972) was used in this study for three reasons. First, this book is more reliable than other biographies about Virginia as the writer is the nephew of Virginia Woolf. Second, the chronology record of this book is complete, from Virginia Woolf's birth (1882) to her death (1912). More importantly, it is a chronology that records the stages of composition of Virginia Woolf's books, her daily life, and especially her illness. We found the psychological states of Virginia Woolf are recorded including *normal*, *recovery*, *restlessness*, *anxiety*, *irritability*, *physical discomfort*, and *mental* *illness*. Thus, we extracted the information on psychological states of Virginia Woolf and the corresponding time for analysis.

Data Search

In this study, we explored some linguistic features as shown in **Table 1**.

We calculated the sentiment and emotion values of each text with Jockers (2015) *syuzhet*, a lexicon-based sentiment analysis package in R (version 3.6.0). To help make the results more reliable, we first calculated the sentiment and emotion values at sentence levels and then computed the mean values of each text as the final value of the text. We chose the *Bing* lexicon (Liu, 2012) to calculate sentiment values and the *NRC* lexicon (Mohammad and Turney, 2013) to calculate the emotion values for the reason that the lexica contain comprehensive lists of sentiment and emotion words and have been widely used in sentiment and emotion research (Calabrese and Cannataro, 2016). With the *Bing* lexicon (Liu, 2012), each sentence was assigned a sentiment value, while each sentence was assigned emotion values from eight perspectives, i.e., anger, anticipation, disgust, fear, joy, sadness, surprise, and trust.

Besides the sentiment and emotion values, we also calculated other linguistic features of each text, such as the number of words, absolutist words, color words, personal pronouns, and question marks. First, following Al-Mosaiwi and Johnstone (2018), we counted the frequency of 19 absolutist words of a text. Second, we counted the frequency of all personal pronouns. It should be noted that we combined the frequency of subject personal pronoun, the object personal pronoun, and the adjective subject pronoun of each personal pronoun of a text (e.g., *i_my_me*) due to the very low occurrence of the pronouns. Third, following Wadsworth et al. (2016), we counted the frequency of each color word of a text, including *red*, *white*, *black*, *green*, *yellow*, *blue*, *purple*, and *gray*. The calculation of all the foregoing linguistic features was performed with a homemade Python script.

Last, due to different text lengths, we normalized the raw frequency of the linguistic features with the following formula.

$$Normalized\ frequency = \frac{Raw\ \ frequency}{Number\ of\ \ words\ in\ the\ text} \times 1,000$$

Data Analysis

The stepwise multiple linear regression analysis was conducted to identify the relation between linguistic features and psychological state of Virginia Woolf, which was implemented with Venables and Ripley's (2002) *MASS*, a powerful package for the statistical and graphical analysis in R (version 3.6.0). To be specific, the independent variables are the linguistics features of texts we extracted. Besides, the dependent variable is the psychological states of Virginia Woolf, which were expressed in numbers for identification and calculation (e.g., 1 represents normal). Our reasons for choosing the stepwise multiple linear regression are listed as follows. First, it identifies the independent variables that can serve as the influential predictors to identify the dependent variable, with the inclusion set at $P < 0.05$. More importantly, it also gains insight into the correlation between the influential predictors and the dependent variable. A positive value of regression coefficient presents a positive relation, and vice versa.

Then, we used the significant linguistic features to build a model for identifying psychological state of Virginia Woolf based on Deep Learning algorithms with the *RapidMiner Studio* (the educational version 9.7). We did so for the following reasons.

First, Deep Learning, as an advanced machine learning algorithm, involves in deciphering the hidden and more complex but meaningful phenomena within the data that the traditional statistical methods struggled to show (Wongso et al., 2017). In simple words, *Deep Learning* may build an automatic but more effective model for identifying psychological state of Virginia Woolf. Second, we can yield an objective evaluated criteria, namely, *accuracy*, with higher values for a better performance (Savoy, 2020).

In conclusion, we determined the significant linguistic features in identifying psychological state of Virginia Woolf, examined their relations, respectively, and built and evaluated the model for identifying the psychological state of Virginia Woolf.

RESULTS

The results show some interesting findings. First, we found the significant linguistic features that jointly predicted the psychological states of Virginia Woolf after examining a total of 65 linguistic features, respectively, in the stepwise multiple regression analysis. The significant linguistic features were the total usage of first-person plural pronouns, i.e., "*we_our_us*," the emotional value of *anger*, and the absolutist word "*everything*." The other linguistic features of this study had no significant effect on predicting the psychological states of Virginia Woolf (*p*-values were larger than 0.05).

Second, we found some correlations between the significant linguistic features and mental health of Virginia Woolf. Specifically, the total use of first-person plural pronouns (Coefficients = −0.007) and the use of emotional value of *anger* (Coefficients = −0.007) were negatively related to mental health of Virginia Woolf. In contrast, the use of the absolutist word "everything" (Coefficients = 0.074) was positively related to mental health of Virginia Woolf.

Last, we used the significant linguistic features to build a model for identifying psychological state of Virginia Woolf, which yielded 86.9% accuracy with Deep Learning algorithms.

DISCUSSION

Several points should be noted based on the results of the present study. First, the total use of first-person plural pronouns correlated negatively with the psychological states of Virginia Woolf. The reasons are listed as follows. First, the use of first-person plural pronouns serves as a mark of social connection, inclusiveness, and belongingness (Pennebaker et al., 2003). Second, fewer first-person plural pronouns serve as evidence of greater self-involvement and selfishness (Pennebaker et al., 2005). Thus, it was the degree of connection and belongingness reflected the psychological states of Virginia Woolf with the total use of first-person plural pronouns. Inconsistent with some previous research, the total use of first-person plural pronouns outperformed than other personal pronouns in predicting psychological states of Virginia Woolf. The result is most likely due to the property of diaries, Virginia Woolf always concerned and recorded her personal activities, thoughts, and feelings. Thus, the use of first-person singular pronouns occurred with a higher frequency but little variation; the use of second and third personal pronouns occurred with a lower frequency; only the use of first-person plural pronouns occurred with a higher frequency and variation. In addition, the total use of first-person plural pronouns highlighted the fluctuations in variation and ensured a significant statistical analysis.

Second, the use of anger words correlated negatively with the psychological states of Virginia Woolf. Many researchers have studied the relation between emotions and psychological states since emotions are one of the key aspects that characterize many mental health conditions (Pennebaker et al., 2003, 2005). However, mixed findings were obtained regarding the relation between emotions and psychological states. For example, Pennebaker et al. (2003) found a linear relation between positive emotion and mental health, and a curvilinear relation between negative emotion and mental health. Some studies found a negative relation between negative emotion and mental health (Kahn et al., 2007; Herbert et al., 2019). They stressed the negative emotions have a detrimental impact on psychological states. Thus, this study complements previous studies, which proved that the anger expression could predict and benefit the psychological states of Virginia Woolf. The reasons are listed as follows. First, anger is an interpersonal feeling, which typically emerges in a close relation (Fehr et al., 1999). Virginia Woolf was involved in a close connection with others and expressed anger, which is beneficial to her psychological health. In addition, angry expression triggers more optimism and shows a higher expectation for the future (Lerner et al., 2003). Last, anger expression helps people realize mistakes and promote problem-solving (Aarts et al., 2010).

Third, the absolutist word "everything" correlated positively with the prediction of the psychological states of Virginia Woolf. Consistent with previous research, absolutist thinking performs an unhealthy and inflexible way of thinking with more absolutist words (Savekar et al., 2019). As expected, the thought patterns of unhealthy psychological individuals are concerned with the feeling of absolutism. It might also signify that absolutist thinking harmed the psychological health of Virginia Woolf. The reasons are listed as follows. First, absolutist thinking disrupts emotion regulation, promotes emotional distress, and make people vulnerable to poor psychological and physical health (Al-Mosaiwi and Johnstone, 2018). Moreover, absolutist people are often perfectionists who see their values, goals, and outcomes as being right (Antoniou et al., 2017). They prefer anger, self-blame, and deprecatory thoughts (Savekar et al., 2019). It is of interest to note that the absolutist word "*everything*" outperformed than other absolutist words we extracted in predicting psychological states of Virginia Woolf. The possible reason is that the absolutist word "*everything*" represents an overly dichotomous thought. The dichotomous thought lacks tolerance and compromise, which may do harm to mental health (Jones et al., 2020). Moreover, "*everything*" served as an indefinite pronoun or premodifier in the diary of Virginia Woolf. Thus, the diachronic change of usage of "*everything*" was more significant than of other absolutist words.

Last, our study verifies the value of the proposed model for predicting the psychological states of Virginia Woolf. Consistent with previous studies, the results proved that a change of psychological states affects the words people used, and the words people used convey a great deal of information about their psychological health. More importantly, our study, similar to ample research, suggested that the model based on linguistic features for predicting psychological states may become increasingly feasible and more accurate. For example, Boukil et al. (2019) proposed an automatic system based on Deep learning algorithm to predict suicidal ideation through analyzing sentiment and feelings expressed on social media. Similarly, Nguyen et al. (2017) extract linguistic features and topics to discriminate depressed online communities from other groups based on machine learning algorithms. Particularly, it yielded 77.6% accuracy in the binary classification of depression vs. Bipolar by Lasso algorithms. Meanwhile, Eichstaedt et al. (2018) built a model with some linguistic predictors to identify depressed *Facebook* users, with a higher prediction accuracy (AUC = 0.72). In addition, de Ávila Berni et al. (2018) developed a model to identify texts proposed by suicidal individuals based on the Naïve-Bayes machine-learning algorithm. The model achieved a higher performance, with an accuracy of 80%. In all, we can develop our proposed model into a system or tool for assessing mental health, monitoring mental disorder, and preventing suicide.

CONCLUSION

In this study, we explored the relation between linguistic features and the psychological states of Virginia Woolf, which generated three findings. First, the result confirmed the total use of first-person plural pronouns, the emotional value of *anger*, and the absolutist word "*everything*" can jointly predict the psychological states of Virginia Woolf. Second, we provided an effective model to predict psychological states of Virginia Woolf. We can further applied the model in various areas such as interventions, therapeutic protocols, and suicide preventions.

Two limitations of this study should be noted. First, we should consider more factors that affect psychological health to improve the explanatory power and usefulness of our equation, such as other linguistic features, social, environmental, economic, and political contexts. Second, we should examine whether our findings are generalizable in any way with more replicative studies on different people, language, and materials.

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/s.

AUTHOR CONTRIBUTIONS

XD: investigation, formal analysis, methodology, validation, and writing–review and editing.

REFERENCES

Aarts, H., Ruys, K. I., Veling, H., Renes, R. A., de Groot, J. H. B., van Nunen, A. M., et al. (2010). The art of anger: reward context turns avoidance responses to anger-related objects into approach. *Psychol. Sci.* 21, 1406–1410. doi: 10.1177/0956797610384152

Al-Mosaiwi, M., and Johnstone, T. (2018). In an absolute state: elevated use of absolutist words is a marker specific to anxiety, depression, and suicidal ideation. *Clin. Psychol. Sci.* 6, 529–542. doi: 10.1177/2167702617747074

Al-Thubaity, A., Alharbi, M., Alqahtani, S., and Aljandal, A. (2018). "A Saudi Dialect Twitter Corpus for Sentiment and Emotion Analysis," in *2018 21st Saudi Computer Society National Computer Conference (NCC)*, (Piscataway: IEEE), 1–6. doi: 10.1109/NCG.2018.8592998

Alvarez-Conrad, J., Zoellner, L. A., and Foa, E. B. (2001). Linguistic predictors of trauma pathology and physical health. *Appl. Cogn. Psychol.* 15, S159–S170. doi: 10.20971/kcpmd.2007.50.2.1

Androutsopoulou, A., Rozou, E., and Vakondiou, M. (2019). Voices of hope and despair: a narrative-dialogical inquiry into the diaries, letters, and suicide notes of Virginia Woolf. *J. Constr. Psychol.* 33, 367–384. doi: 10.1080/10720537.2019.1615015

Antoniou, E. E., Bongers, P., and Jansen, A. (2017). The mediating role of dichotomous thinking and emotional eating in the relationship between depression and BMI. *Eating Behav.* 26, 55–60. doi: 10.1016/j.eatbeh.2017.01.007

Barnes, D. H., Lawal-Solarin, F. W., and Lester, D. (2007). Letters from a suicide. *Death Stud.* 31, 671–678. doi: 10.1080/07481180701405212

Bell, A. O. (1984). *The Diary of Virginia Woolf (1936-1941)*. San Diego: Harvest.

Bell, Q. (1972). *Virginia Woolf: A Biography*. New York: Harcourt Brace Jovanovich.

Blodgett, H. (1989). A woman writer's diary: Virginia woolf revisited. *Prose Stud.* 12, 57–71. doi: 10.1080/01440358908586360

Boals, A., and Klein, K. (2005). Word use in emotional narratives about failed romantic relationships and subsequent mental health. *J. Lang. Soc. Psychol.* 24, 252–268. doi: 10.1177/0261927X05278386

Boeira, M. V., Berni, G. Á, Passos, I. C., Kauer-Sant'Anna, M., and Kapczinski, F. (2016). Virginia Woolf, neuroprogression, and bipolar disorder. *Braz. J. Psychiatry* 39, 69–71. doi: 10.1590/1516-4446-2016-1962

Boukil, S., El Adnani, F., Cherrat, L., El Moutaouakkil, A. E., and Ezziyyani, M. (2019). "Deep learning algorithm for suicide sentiment prediction," in *Advanced Intelligent Systems for Sustainable Development (AI2SD'2018)*, ed. M. Ezziyyani (Berlin: Springer International Publishing), doi: 10.1007/978-3-030-11884-6_24

Briggs, K. (2011). The making of A Writer's Diary: being extracts from the diary of Virginia Woolf. *Textual Pract.* 25, 1033–1050. doi: 10.1080/0950236x.2011.618459

Calabrese, B., and Cannataro, M. (2016). "Sentiment analysis and affective computing: Methods and applications," in *Brain-Inspired Computing*, eds K. Amunts, L. Grandinetti, T. Lippert, and N. Petkov (Berlin: Springer International Publishing).

Cambria, E. (2016). Affective computing and sentiment analysis. *IEEE Intell. Syst.* 31, 102–107. doi: 10.1109/MIS.2016.31

Campbell, R. S., and Pennebaker, J. W. (2003). The secret life of pronouns flexibility in writing style and physical health. *Psychol. Sci.* 14, 60–65. doi: 10.1111/1467-9280.01419

Chatterjee, A., Gupta, U., Chinnakotla, M. K., Srikanth, R., Galley, M., and Agrawal, P. (2019). Understanding Emotions in Text Using Deep Learning and Big Data. *Comput. Hum. Behav.* 93, 309–317. doi: 10.1016/j.chb.2018.12.029

Ciullo, F., Zucco, C., Calabrese, B., Agapito, G., Guzzi, P. H., and Cannataro, M. (2016). "Computational challenges for sentiment analysis in life sciences," in *2016 International Conference on High Performance Computing & Simulation (HPCS)*, (Innsbruck: IEEE), 419–426. doi: 10.1109/HPCSim.2016.7568365

Cohn, M. A., Mehl, M. R., and Pennebaker, J. W. (2004). Linguistic markers of psychological change surrounding September 11, 2001. *Psychol. Sci.* 15, 687–693.

D'Andrea, A., Ferri, F., Grifoni, P., and Guzzo, T. (2015). Approaches, tools and applications for sentiment analysis implementation. *Int. J. Comput. Appl.* 125, 26–33. doi: 10.5120/ijca2015905866

de Ávila Berni, G., Rabelo-da-Ponte, F. D., Librenza-Garcia, D., Boeira, M. V., Kauer-Sant'Anna, M., Passos, I. C., et al. (2018). Potential use of text classification tools as signatures of suicidal behavior: a proof-of-concept study using Virginia Woolf's personal writings. *PLoS One* 13:e0204820. doi: 10.1371/journal.pone.0204820

Desmet, B., and Hoste, V. (2013). Emotion detection in suicide notes. *Expert Syst. Appl.* 40, 6351–6358. doi: 10.1016/j.eswa.2013.05.050

Eichstaedt, J. C., Smith, R. J., Merchant, R. M., Ungar, L. H., Crutchley, P., Preoţiuc-Pietro, D., et al. (2018). Facebook language predicts depression in medical records. *Proc. Natl. Acad. Sci.* 115, 11203–11208. doi: 10.1073/pnas.1802331115

Ekman, P. (1993). Facial expression and emotion. *Am. Psychol.* 48, 384–392.

Fehr, B., Baldwin, M., Collins, L., Patterson, S., and Benditt, R. (1999). Anger in close relationships: an interpersonal script analysis. *Pers. Soc. Psychol. Bull.* 25, 299–312. doi: 10.1177/0146167299025003003

Garcia, D., Garas, A., and Schweitzer, F. (2012). Positive words carry less information than negative words. *EPJ Data Sci.* 1:3. doi: 10.1140/epjds3

Graves, K. D., Schmidt, J. E., Bollmer, J., Fejfar, M., Langer, S., Blonder, L. X., et al. (2005). Emotional expression and emotional recognition in breast cancer survivors: a controlled comparison. *Psychol. Health* 20, 579–595. doi: 10.1080/0887044042000334742

Herbert, C., Bendig, E., and Rojas, R. (2019). My sadness – our happiness: writing about positive, negative, and neutral autobiographical life events reveals linguistic markers of self-positivity and individual well-being. *Front. Psychol.* 9:2522. doi: 10.3389/fpsyg.2018.02522

Jockers, M. (2015). *Syuzhet: Extract Sentiment and Plot Arcs from Text*. Available online at https://github.com/mjockers/syuzhet. (accessed on Apr 12, 2021).

Jones, L. S., Anderson, E., Loades, M., Barnes, R., and Crawley, E. (2020). Can linguistic analysis be used to identify whether adolescents with a chronic illness are depressed? *Clin. Psychol. Psychother.* 27, 179–192. doi: 10.1002/cpp.2417

Kahn, J. H., Tobin, R. M., Massey, A. E., and Anderson, J. A. (2007). Measuring emotional expression with the linguistic inquiry and word count. *Am. J. Psychol.* 120:263. doi: 10.2307/20445398

Kim, K., Choi, S., Lee, J., and Sea, J. (2019). Differences in linguistic and psychological characteristics between suicide notes and diaries. *J. Gen. Psychol.* 146, 391–416. doi: 10.1080/00221309.2019.1590304

Lei, L., and Liu, D. (2021). *Conducting Sentiment Analysis*, 1st Edn. Cambridge: Cambridge University Press.

Lennox, R. J., Veríssimo, D., Twardek, W. M., Davis, C. R., and Jarić, I. (2020). Sentiment analysis as a measure of conservation culture in scientific literature. *Conserv. Biol.* 34, 462–471. doi: 10.1111/cobi.13404

Lerner, J. S., Gonzalez, R. M., Small, D. A., and Fischhoff, B. (2003). Effects of fear and anger on perceived risks of terrorism: a national field experiment. *Psychol. Sci.* 14, 144–150. doi: 10.1111/1467-9280.01433

Liu, B. (2012). Sentiment analysis and opinion mining. *Synth. Lect. Hum. Lang. Technol.* 5, 1–167.

Mohammad, S. M., and Turney, P. D. (2013). Crowdsourcing a word–emotion association lexicon. *Comput. Intell.* 29, 436–465. doi: 10.1111/j.1467-8640.2012.00460.x

Molina Beltrán, C., Segura Navarrete, A. A., Vidal-Castro, C., Rubio-Manzano, C., and Martínez-Araneda, C. (2019). Improving the affective analysis in texts: automatic method to detect affective intensity in lexicons based on Plutchik's wheel of emotions. *Electron. Libr.* 37, 984–1006. doi: 10.1108/EL-11-2018-0219

Moreno-Blanco, D., Ochoa-Ferreras, B., Gárate, F. J., Solana-Sánchez, J., Sánchez-González, P., and Gómez, E. J. (2020). "Evaluation and Comparison of Text Classifiers to Develop a Depression Detection Service," in *XV Mediterranean Conference on Medical and Biological Engineering and Computing – MEDICON 2019*, eds J. Henriques, N. Neves, and P. de Carvalho (Berlin: Springer International Publishing), doi: 10.1007/978-3-030-31635-8_146

Newman, M. L., Pennebaker, J. W., Berry, D. S., and Richards, J. M. (2003). Lying words: predicting deception from linguistic styles. *Pers. Soc. Psychol. Bull.* 29, 665–675. doi: 10.1177/0146167203029005010

Nguyen, T., O'Dea, B., Larsen, M., Phung, D., Venkatesh, S., and Christensen, H. (2017). Using linguistic and topic analysis to classify sub-groups of online depression communities. *Multimed. Tools Appl.* 76, 10653–10676. doi: 10.1007/s11042-015-3128-x

Papapicco, C., and Mininni, G. (2020). Impact memes: PhDs HuMor(e). *Multimed. Tools Appl.* 79, 35973–35994. doi: 10.1007/s11042-020-09166-0

Pennebaker, J. W., Mehl, M. R., and Niederhoffer, K. G. (2003). Psychological aspects of natural language use: our words, our selves. *Ann. Rev. Psychol.* 54, 547–577. doi: 10.1146/annurev.psych.54.101601.145041

Pennebaker, J. W., Slatcher, R. B., and Chung, C. K. (2005). Linguistic markers of psychological state through media interviews: John Kerry and John Edwards in 2004, Al Gore in 2000. *Anal. Soc. Issues Public Policy* 5, 197–204. doi: 10.1111/j.1530-2415.2005.00065.x

Plutchik, R. (1980). "Theories of emotion," in *Emotion: Theory, Research and Experiences*, eds R. Plutchik and H. Kellerman (New York, NY: Academic Press).

Rendalkar, S., and Chandankhede, C. (2018). "Sarcasm detection of online comments using emotion detection," in *2018 International Conference on Inventive Research in Computing Applications (ICIRCA)*, (Piscataway: IEEE), 1244–1249. doi: 10.1109/ICIRCA.2018.8597368

Rodrigues, R. G., das Dores, R. M., Camilo-Junior, C. G., and Rosa, T. C. (2016). SentiHealth-Cancer: a sentiment analysis tool to help detecting mood of patients in online social networks. *Int. J. Med. Inform.* 85, 80–95. doi: 10.1016/j.ijmedinf.2015.09.007

Rude, S., Gortner, E.-M., and Pennebaker, J. (2004). Language use of depressed and depression-vulnerable college students. *Cogn. Emot.* 18, 1121–1133. doi: 10.1016/j.beth.2006.01.004

Savekar, A., Tarai, S., and Singh, M. (2019). "Linguistic markers in individuals with symptoms of depression in bi-multilingual context," in *Early Detection of Neurological Disorders Using Machine Learning Systems*, eds S. Paul, P. Bhattacharya, and A. Bit (Pennsylvania: IGI Global), 216–240. doi: 10.4018/978-1-5225-8567-1.ch012

Savoy, J. (2020). *Machine Learning Methods for Stylometry: Authorship Attribution and Author Profiling*. Berlin: Springer International Publishing.

Simmons, R. A., Chambless, D. L., and Gordon, P. C. (2008). How do hostile and emotionally overinvolved relatives view relationships?: what relatives' pronoun use tells us. *Fam. Process* 47, 405–419. doi: 10.1111/j.1545-5300.2008.00261

Tausczik, Y. R., and Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *J. Lang. Soc. Psychol.* 29, 24–54. doi: 10.1177/0261927X09351676

Venables, W. N., and Ripley, B. D. (2002). *Modern Applied Statistics with S*. New York: Springer.

Wadsworth, F. B., Vasseur, J., and Damby, D. E. (2016). Evolution of vocabulary in the poetry of sylvia plath. *Digit. Scholarsh. Humanit.* 32, 660–671.

Wongso, R., Luwinda, F. A., Trisnajaya, B. C., Rusli, O., and Rudy. (2017). News Article Text Classification in Indonesian Language. *Procedia Comput. Sci.* 116, 137–143. doi: 10.1016/j.procs.2017.10.039

Ziemer, K. S., and Korkmaz, G. (2017). Using text to predict psychological and physical health: a comparison of human raters and computerized text analysis. *Comput. Hum. Behav.* 76, 122–127. doi: 10.1016/j.chb.2017.06.038

Zucco, C., Calabrese, B., Agapito, G., Guzzi, P. H., and Cannataro, M. (2020). Sentiment analysis for mining texts and social networks data: methods and tools. *WIREs Data Min. Knowl. Discov.* 10:e1333. doi: 10.1002/widm.1333

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.

frontiers | Frontiers in Psychology

REVIEW
published: 07 July 2022
doi: 10.3389/fpsyg.2022.911068

Check for updates

How Emotion Relates to Language and Cognition, Seen Through the Lens of Evaluative Priming Paradigms

Michaela Rohr and Dirk Wentura*

Cognitive Psychology Department, Saarland University, Saarbrücken, Germany

Affect and emotion are essential aspects of human life. These states or feelings signal personally relevant things or situations and color our memories and thoughts. Within the area of affective or emotion processing, evaluation–the assessment of the valence associated with a stimulus or event (i.e., its positivity or negativity)–is considered a fundamental process, representing an early and crucial stage in constructivist emotion theories. Valence evaluation is assumed to occur automatically when encountering a stimulus. But does this really apply always, even if we simply see a word? And if so, what exactly is processed or activated in memory? One approach to investigating this evaluative process uses behavioral priming paradigms and, first and foremost, the evaluative priming paradigm and its variants. In the present review, we delineate the insights gained from this paradigm about the relation of affect and emotion to cognition and language. Specifically, we reviewed the empirical evidence base with regard to this issue as well as the proposed theoretical models of valence evaluation, specifically with regard to the nature of the representations activated *via* such paradigms. It will become clear that affect and emotion are foremost (and, perhaps, even exclusively) triggered by evaluative priming paradigms in the sense that semantic affective knowledge is activated. This knowledge should be modeled as being active in working memory rather than in long-term memory as was assumed in former models. The emerging evidence concerning the processing of more specific emotion aspects gives rise to the assumption that the activation of these semantic aspects is related to their social importance. In that sense, the fast and (conditionally) automatic activation of valence and other emotion aspects in evaluative priming paradigms reveals something about affect and emotion: Valence and specific emotion aspects are so important for our daily life that encountering almost any stimulus entails the automatic activation of the associated valence and other emotion aspects in memory, when the context requires it.

Keywords: evaluative priming, affect, emotion, semantic representation, valence

OPEN ACCESS

Edited by:
Maya Michelle Khanna,
Creighton University, United States

Reviewed by:
Sascha Topolinski,
University of Cologne, Germany
Kerstin Fröber,
University of Regensburg, Germany

****Correspondence:***
Michaela Rohr
m.rohr@mx.uni-saarland.de

Specialty section:
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

Received: *01 April 2022*
Accepted: *07 June 2022*
Published: *07 July 2022*

Citation:
Rohr M and Wentura D (2022) How Emotion Relates to Language and Cognition, Seen Through the Lens of Evaluative Priming Paradigms. Front. Psychol. 13:911068. doi: 10.3389/fpsyg.2022.911068

INTRODUCTION

Affect and emotion are essential aspects of human life. These states or feelings signal personally relevant things or situations and color our memories and thoughts. Within the area of affective or emotion processing, evaluation–the assessment of the valence associated with a stimulus or event (i.e., its positivity or negativity) –is considered a fundamental process, representing an early and crucial stage in constructivist emotion theories. Valence evaluation is assumed to occur automatically when encountering a stimulus. But does this really apply always even if we simply see a word? If so, what exactly is processed or activated in memory? One approach to investigating this evaluative process uses behavioral priming paradigms, and, first and foremost, the evaluative priming paradigm and its variants.

In the present review, we delineate the insights gained from this paradigm about the relation of affect and emotion to cognition and language. We will start with a theoretical outline of the concepts of emotion, affect, and evaluation to avoid misunderstandings, because researchers do not completely agree upon these concepts' definitions (see, e.g., Moors and Scherer, 2013). Of importance for the present review, this delineation shapes our view on how the processing of stimuli in evaluative priming paradigms relates to emotion and affect.

We then introduce the reader to the evaluative priming paradigm and its variants [see Herring et al. (2013) for a meta-analysis]. In the basic version of the evaluative priming paradigm, target items (e.g., words) are categorized according to their valence (Fazio et al., 1986). The target is preceded by a briefly presented prime stimulus whose valence is either congruent to the target's response category or not. Typically, faster responses and/or fewer errors are observed in cases of congruency, which is taken as an indication of automatic evaluation of the prime. We outlined the empirical evidence gathered with the paradigm, focusing on three specific questions: (1) Are "hot" emotion-related processes involved in the paradigm? Or do effects rely on the "cold" semantic categorization process (es) (or do both play a role)? We used the term "hot" to refer to all aspects related to emotion or affect that cannot be explained simply by pure semantic processing, such as, for example, the involvement of facial muscle responses, or effects relating to embodiment (e.g., modality switch costs). After having answered this question, we will address the more specific question: (2) How is valence mentally represented according to the evidence accrued with the evaluative priming paradigm? To answer this question, we describe the different task contexts under which effects in the evaluative priming paradigm have been observed (or not). It will become clear that the assessment of valence is not an integral part of prime word processing under all conditions despite its functional relevance for adaptive behavior. That is, valence assessment is not automatic in that it does not unconditionally meet all automaticity criteria (i.e., being fast, efficient, unintentional, non-conscious, and independent of current goals; Moors, 2007). It will become clear that, despite three decades of research, the existing theoretical models cannot explain all observed effects. Thus, we will propose a working memory account, integrating existing mechanisms into a coherent picture of evaluative priming. We will then ask a final question: (3) Beyond valence, can further emotion aspects—such as relevance, potency, or the specific emotion category—be involuntarily activated (and what insights may the answer to this question provide)? As the reader will see, the existing evidence is not easily explainable by (pure) semantic processing accounts so that, indeed, the evaluative priming paradigm and its variants reveal something about affect and emotion.

WHAT EVALUATIVE PRIMING RESEARCHERS MEAN WHEN THEY TALK ABOUT AFFECT AND EMOTION

Affective phenomena are a very popular and widely researched topic (Dukes et al., 2021), and priming paradigms are only one approach to gain insights into them. To prevent issues akin to the parable of the blind men describing an elephant, we first provide definitions of core concepts: Affect is usually defined in a very general sense as "any experience of feeling or emotion… ranging from the simplest to the most complex sensations" (VandenBos, 2007, p. 26). Next to cognition and conation, it is identified as one of the essential components of the mind. Affect can range from a vague feeling, not bound to a specific object and, simply, phenomenally experienced to a specific emotional state that can involve cognitive aspects and can be reflected on (e.g., I experience disappointment for not having achieved a self-set goal). Often, however, researchers use the word affect in a narrower sense when referring to coarse, unspecific states that can be described as positive and negative, in contrast to specific emotional states. An emotion is seen as a temporary, dynamic process encompassing "interrelated, synchronized changes" in several components, including cognition, neurophysiology, motor expression, motivation, and subjective feeling "in response to the evaluation of an internal or external stimulus event as relevant to major concerns of the organism" (Scherer, 2005, p. 313). Important differences from affect lie in the circumscribed phenomenal feelings and related cognitions elicited by a specific stimulus or event.

For both concepts, the process of evaluation or appraisal plays a primary role, that is, the process by which an individual assigns subjective, affective meaning to stimuli. Evaluation can be broadly defined as the outcome of such a judgment process (e.g., Cunningham et al., 2007), no matter how many and which processing steps are exactly involved. However, evaluation (especially in the context of evaluative priming paradigms) is typically defined more narrowly as a mechanism that automatically evaluates all incoming stimulus information as "good/positive/pleasant" or "bad/negative/unpleasant." More precisely, the stored mnemonic representation of the evaluative aspect of a stimulus is assumed to be activated along with its semantic representation.[1] Whereas, the term evaluation

[1]Of course, we do not assume that every mental concept has an existing evaluative connotation. When encountering a stimulus for the first time, evaluation and/or an affective response occur for the first time, and the response is constructed online. By and large, empirical research employing the evaluative priming paradigm has, however, focused on clearly evaluative (mostly normative) stimuli whose evaluative connotation can be assumed to be stored in memory.

emphasizes the cognition involved in the judgment, the term appraisal places more weight on different aspects of (and thus different kinds of information involved in) emotional processing (i.e., physiological state, expression, action tendencies, cognition, and feeling; Scherer and Moors, 2019).

This already makes clear that researchers with different backgrounds and research interests have different perspectives on the processes that assign subjective, affective meaning to stimuli. Some take a more cognitive stance in examining relatively stable affective phenomena, such as, for example, attitudes (Cunningham et al., 2007). Others focus more on the physiological and phenomenal aspects of affect and emotion and the temporary nature of emotional phenomena (Scherer and Moors, 2019). Evaluative priming research falls on the more cognitive side, as will become clear shortly. Thus, one might ask in what sense, if any, is evaluative priming related to affect and emotion? One general response is that emotion and cognition are deeply intertwined (see, e.g., Moors, 2007; Storbeck and Clore, 2007; Barrett, 2014), and evaluative priming research can thus, indeed, answer questions related to emotion and affect: When we extract information from our surroundings, translating it into "meaning," we do so based on cognitive processes. This process of meaning making implies that we do not take the sensory input as it is, but we translate it into something. As Moors (2007, p. 1238) puts it: cognitive processes are those that "mediate between variable input-output relations by means of representations." She proposes that "the unique feature of emotion has to do with the content of the representations involved in the transition from stimulus input to emotion." This statement makes the relation of emotion to cognition as well as the importance of language for this relation very clear: Emotional information–as any other type of information–is processed *via* extracting some sort of category[2]. and mapping it onto a mental (potentially modal, embodied) representation, but the type of mental representation or the extracted information encompasses aspects specific to emotion and affect. Language, in turn, can be used to label the representation, and the representation might be partly or completely semantic in nature.

Two further distinctions are worth making, especially with regard to linguistic stimuli: the first is the one between *denotative* and *connotative* meaning (e.g., Osgood et al., 1957). The *denotative* meaning refers to the semantic mental representation of a word: For example, when reading the word HOUSE, a semantic mental representation is activated, which refers to the real object being referred to. The denotative meaning is identical or similar across individuals and the basis for communicating about the referred-to objects. *Connotative* aspects, by contrast, refer to all subjective, affective, or emotional associations of a stimulus (e.g., one could think of a warm and cozy home when reading HOUSE, or an old, gray, run-down tenement). These associations depend on an individual's stored mental representations created across their lifespan; thus, connotative aspects are multidimensional, and their content varies strongly between individuals (Osgood et al., 1957). Thus, when studying evaluation using priming paradigms, it is often the connotative aspect of the stimuli that is of interest. Sometimes, however, this distinction is not considered (e.g., when using emotional facial expressions, researchers often focus on their denotative meaning as positive or negative or representing a specific emotion). The second important distinction is the one between *affective* and *semantic valence* (Itkes and Kron, 2019). Itkes and Kron correctly pointed out that researchers often do not sufficiently differentiate between affective phenomena that are part of an emotional response–in the sense that they involve a physiological response or a potentially fleeting change in emotional feeling (i.e., affective valence, "hot" affect) –and affective phenomena that merely involve activation of factual knowledge (i.e., memory representations) of the valence associated with an object or event (i.e., semantic valence, "cold" affect). This distinction is important for evaluative priming, as there is ongoing debate about what exactly is activated in this paradigm and what mechanisms underlie the observed effects. We will come back to this issue later on.

Studying valence evaluation with priming paradigms is, therefore, a research field at the intersection of affect, cognition, and language: It uses predominantly linguistic stimuli (but also facial expressions, pictures, sounds) to study the automatic (i.e., fast, efficient, non-conscious, goal-independent, uncontrollable; Moors and De Houwer, 2006) assessment of valence, including the underlying processes, processing steps, or mechanisms, and mental representations.

IS SEMANTIC OR AFFECTIVE VALENCE ACTIVATED IN EVALUATIVE PRIMING?

A core question in this area of research has been whether the affective connotation of words (and other stimuli) renders evaluative priming different from other forms of semantic processing, and whether the underlying processes or the activated mental representations differ from other semantic processes and representations. More specifically, researchers are interested in whether evaluative priming involves "hot" affective aspects (i.e., affective valence in terms of Itkes and Kron, 2019) –that is, activation of emotion components, such as facial expressions or physiological changes. We will now summarize the insights that empirical research has provided so far.

The question of whether evaluative priming differs from other forms of semantic categorization has been posed by this research field since its beginning (e.g., Fazio et al., 1986; Klauer and Musch, 2003; Storbeck and Robinson, 2004); however, only few empirical studies have directly studied this issue. One focus has been on whether evaluative categorization is more obligatory than semantic categorization and whether it occurs before or independently from semantic analysis (e.g., Zajonc, 1980; Nummenmaa et al., 2010). While many studies have

[2]We use the term category here in a cognitive-science sense. As Harnad (2005) (p. 20) puts it: "We, organisms, are sensorimotor systems. Things in the world come in contact with our sensory surfaces, and we interact with them based on what that sensorimotor contact "affords." All of our categories consist of ways we behave differently toward different *kinds* of things – things we do or do not eat, mate with, or flee from; or the things that we describe, through our language, as prime numbers, affordances, absolute discriminables, or truths." From that viewpoint, emotions can be seen as one kind of category (see also Boster, 2005).

investigated this question using categorization tasks without primes (e.g., Nummenmaa et al., 2010; Lai et al., 2012), some have employed an evaluative priming paradigm[3] These early studies, by and large, showed comparable results: If evaluative categorization was the task at hand, evaluative priming effects emerged; if semantic categorization (e.g., person vs. animal) was the task at hand, semantic priming effects emerged (Kemp-Wheeler and Hill, 1992; Klinger et al., 2000; De Houwer and Randell, 2002; De Houwer et al., 2002), with no evidence for the additional activation of valence (see also Section How Is Valence Represented in Conceptual Representation Systems?). There is, however, one notable exception: Storbeck and Robinson (2004) systematically compared semantic categorization to evaluative categorization using the same stimuli, which varied not only with regard to valence but also with regard to category membership (i.e., animals vs. texture-related concepts vs. religious concepts; words in Exp. 1-3, pictures in Exp. 4-5). In four experiments, category congruency effects emerged and no evaluative priming effects, regardless of whether the task was semantic or evaluative categorization. Evaluative priming effects emerged only in one experiment, in which there was no variation of semantic category (Exp. 3). The authors explained this somewhat surprising finding with the salient variation in semantic category membership. According to them, evaluative priming effects only emerge when there is no (close) semantic relationship between primes and targets (but see Section How Is Valence Represented in Conceptual Representation Systems?), or when there is little semantic variation in a given task context. This already pointed to the conditionality of evaluative priming effects, an aspect that has been more systematically investigated in recent years and which we will elaborate on below. Of importance for the affective vs. semantic valence issue here is that there is evidence that evaluative categorization does not occur earlier than or independently of semantic categorization, but that semantic categorization seems to be the default processing operation when encountering (a wide range of) stimuli. Evaluative priming effects occur only when evaluation is important in the given context (e.g., due to the task set or attentional focus; see below), which is at odds with affective primacy.

Notwithstanding this documented predominance of semantic processes in the paradigm, one can ask whether a "hot" affect-related response (i.e., a physiological affective response or a partial reinstatement of such, as assumed by embodiment perspectives) is triggered by the prime, in addition to some semantic activation, and whether an affective state can influence evaluative priming effects.

With regard to the latter question, there are some hints that an affective state influences the effects: Foroni and Semin (2012) showed that inhibition of the zygomaticus major muscle–the muscle that pulls the lip corners apart in smiling–prevents evaluative priming effects. They argued that this is due to the involvement of affect-related, embodied processes in the paradigm. An alternative explanation would be that the muscle inhibition reduces positive affect. In support of this, Storbeck and Clore (2008) showed that participants in a neutral or happy mood showed evaluative priming effects but participants in a sad mood did not. These authors argued that affect influences the accessibility and use of semantic associations in memory, and that negative affect can inhibit activation spread in semantic networks. Thus, the observed effect would be the result of an affect-cognition interaction, and not evidence for the involvement of affect in evaluative priming *per se*. In a similar vein, Lemonnier and Alexopoulos (2019) examined the modulation of the evaluative priming effect by brief phasic affect, that is, trial-by-trial affective variations induced by positive and negative music or the manipulation of proprioceptive facial feedback *via* different facial muscle postures. Conceptually, they replicated Storbeck and Clore's (2008) results: Evaluative priming was observed in the neutral and positive conditions, but not in the negative affect conditions (also see Topolinski and Deutsch, 2013, for an influence of phasic affect on semantic priming). Consistent with this, De Saedeleer and Pourtois (2016; Exp.1) found that trait worry weakens the evaluative priming effect. Taken together, this line of research clearly shows that "hot" affect can influence evaluative priming effects; however, it seems unlikely that the processes underlying evaluative priming themselves encompass "hot" affect-related aspects, even though the available evidence cannot rule this out.

Another approach to the topic is the investigation of arousal influences on evaluative priming. Arousal is defined as the degree of physiological activation, and it is considered the second important dimension of the (semantic) affective space (Osgood et al., 1957; Russell and Barrett, 1999). Thus, if arousal affected evaluative priming effects, "hot" affect-related aspects might play a role in the paradigm. In three studies, Herring et al. (2015) used affective images rated low vs. high on arousal to examine whether prime arousal has an influence on evaluative priming effects. In their studies, greater prime arousal increased evaluative priming if primes were presented parafoveally, but not when primes were presented foveally. The authors argued this effect resulted from greater competition between sensory and response processes for high arousing parafoveal stimuli, which slowed down target processing in incongruent trials. Taking a more general perspective, the effects can be explained with arousal-biased competition theory (Mather and Sutherland, 2011), which proposes that arousal increases activation of salient or relevant stimuli and dampens the activation of less salient competitors. Consequently, the effects of Herring et al. can be explained by arousal-induced changes in attention and early vision, which are of particular relevance for parafoveal processing. If attentional processes are of less importance, as for example, in central vision, then arousal does not seem to play a role for evaluative priming. In line with this, Hinojosa et al. (2009) found that arousal did not influence reaction times in evaluative priming, although event-related potentials showed clear processing of

[3] We focus here on the evaluative categorization task with positive/negative primes and targets. There is also a very similar paradigm: the affect misattribution procedure (Payne et al., 2005; note this was, initially, also referred to as the affective priming paradigm, Murphy and Zajonc, 1993). Reviewing the evidence from this paradigm is beyond the scope of the present review. In a nutshell, research using this paradigm provides different answers to the questions outlined here; the underlying processes and mechanisms are different (see, e.g., Gawronski et al., 2010 see Payne and Lundberg, 2014, for an overview), as are the underlying activated representations (e.g.,De Houwer and Tucker Smith, 2013; Rohr et al., 2015).

the arousal aspect. However, Zhang et al. (2012) did observe a behavioral influence of valence and arousal: Using picture primes and word targets, they found that valence congruency effects were enhanced by high prime arousal. They also reported an effect of arousal congruency, which, however, was not mirrored in the corresponding event-related potentials. Also, as they used Chinese words (i.e., words of a logographic language), it is unclear whether the observed effects generalize to phonographic languages as well. So far, the very few existing studies suggest that arousal can impact evaluative priming effects; however, this influence seems to arise from arousal-induced influences on attention and downstream processing, not from the involvement of "hot" affect-related aspects in evaluative priming.

Thus, taken together, the existing evidence speaks against an involvement of affective valence in the evaluative priming paradigm. However, it remains a possibility that some "hot" aspects can be involved in the paradigm–as suggested by embodiment perspectives (e.g., Niedenthal et al., 2003), –and that research simply has not yet found the right approach to capture this influence (e.g., see Itkes et al., 2017, for an intriguing test of affective vs. semantic valence in the affective Simon paradigm).

HOW IS VALENCE REPRESENTED IN CONCEPTUAL REPRESENTATION SYSTEMS?

The Evaluative Priming Paradigm and Its Mechanisms

Given the evidence presented above, it seems obvious that some form of semantic mental representation becomes activated in evaluative priming paradigms. But how is the activated valence mentally represented? This question cannot be answered without a close look at the mechanisms involved in the paradigm: Evaluative priming with the evaluative decision task is a variant of response interference tasks, like Stroop or flanker tasks (see Klauer et al., 1997; Wentura, 1999; also Wentura and Degner, 2010b). In the classical setup, target stimuli (e.g., words, pictures, faces, sounds) are categorized according to their valence (i.e., as positive or negative). The target (e.g., the word LOVE) is preceded by a briefly presented prime stimulus (e.g., the word BIRTHDAY) whose valence is either congruent to the target's response category or not. Typically, faster responses and/or fewer errors are observed in cases of congruency, which is taken as an indication of involuntary evaluation of the prime; that is, it is assumed that the valence of the task-irrelevant prime stimulus is automatically assessed. Thus, the paradigm is a parsimonious way to examine the automatic evaluation of the presented prime stimulus (in the example above: its connotative meaning) and the underlying processes and specific facilitatory conditions; it can potentially also reveal something about affect-cognition interactions. Moreover, given the task's setup, the question under investigation can be reframed as "Is the prime (distractor) compatible or incompatible with the response required by the target?" From this perspective, the most parsimonious explanation of the observed effects is based on response compatibility, that is, the assumption that the prime can facilitate either the correct or the incorrect response, with ensuing effects on response times [stimulus-response (or S-R)-based evaluative priming; see De Houwer, 2003]. Thus, evaluative priming can be conceptualized as just another variant of semantic categorization with response priming as the underlying mechanism [e.g., see Dehaene et al., 1998; Kunde et al., 2003; Rohr and Wagner, 2020 for studies with numbers; Klinger et al., 2000; Kiesel et al., 2006 for studies using other semantic categories]. Put differently: In the evaluative categorization task, stimulus valence is processed because it is response relevant. Thus, valence activated by a prime could just be seen as any other (response-relevant) semantic category. As such, a priming effect reveals (goal-dependent) involuntary activation of the prime's valence, but it does not reveal anything about the links between positive/negative concepts in semantic memory, for example, and thus provides little information about how valence is mentally represented.

However, the evaluative priming paradigm can be turned into a variant of the associative/semantic priming paradigm by changing task instructions (e.g., see Neely, 1991; McNamara, 2005; Wentura and Degner, 2010b). Concretely, if the task is to decide whether a target is a word or non-word, or to simply pronounce the target as quickly as possible, then response compatibility can no longer be the underlying mechanism. For example, a pronunciation task will mix valence-congruent prime/target pairs (e.g., *FLOWER-CAKE*) and valence-incongruent recombinations of prime/target pairs (e.g., *WAR-SUNSHINE*). Any resulting priming effects cannot be explained easily by S-R-based processes, but explanation must take into account processes that focus on the mental representation of prime and target and how activation of the prime representation helps or hinders access to the target representation [i.e., stimulus-stimulus (or S-S)-based priming; De Houwer, 2003].

Thus, changing an evaluative priming task from evaluative judgment to, for example, pronunciation is *prima facie* associated with a change in the underlying mechanism from response interference to encoding facilitation. This change brings up a very intriguing question: Will any prime render all stimuli of the same valence more accessible, such that a target of the same valence can be processed more quickly? Please note that this "encoding facilitation hypothesis" (Spruyt et al., 2002) with respect to valence as a concept differs from encoding facilitation for other semantic categories simply because (almost) every stimulus is assumed to have an evaluative connotation. Thus, the semantic "category" of positive/negative valence is much broader than other semantic categories. If this assumption holds, it suggests that valence is not only an important, easy-to-retrieve feature in the representation of an individual concept, but a powerful "connector" between representations.

We will now review the evidence in favor and against this encoding facilitation hypothesis, because – as mentioned earlier – this evidence allows for more insights into the underlying mental representations. That is, in the following paragraphs, we focus on evaluative priming evidence obtained with tasks other than the evaluative decision task. For a general overview, the reader is referred to Klauer and Musch (2003) and the meta-analysis of Herring et al. (2013). Having reviewed this evidence, we will outline the theoretical models that have been developed based on this research.

The Encoding Facilitation Hypothesis

Early studies [Hermans et al. (1994, Exp. 2); Bargh et al. (1996), three experiments] provided initial evidence for evaluative priming with the pronunciation task, and thus for the assumption that evaluative concepts might facilitate access to other concepts of the same valence. Further evidence was found by Duckworth et al. (2002) in two experiments and Hermans et al. (2001, Exp. 2). However, later studies failed to replicate these early findings. For example, Klauer and Musch (2001) conducted a total of five experiments, all with null results (see also Hermans et al., 1996, Exp. 8; De Houwer et al., 1998; De Houwer and Hermans, 1999; Spruyt et al., 2004). Notably, in six experiments, Glaser and Banaji (1999) even found robust *reversed* effects, that is, faster responses in incongruent trials than congruent ones (see also Maier et al., 2003; Berner and Maier, 2004). Thus, early research using the pronunciation task yielded inconsistent results and provided no clear picture on whether stimulus valence can be activated when it is not task relevant. Using semantic classification of non-valence categories [i.e., animal/person (Exp. 1), object/person (Exp. 2)], De Houwer and Randell (2002) failed to obtain evidence for evaluative priming. Similarly, Klauer and Musch (2002) (Exp. 1–4) used four different binary classifications orthogonal to valence; none of the experiments showed an evaluative priming effect. Likewise, Hermans et al. (1998), as well as Rothermund and Wentura (1998), found no results with a color-naming task (see Warren, 1972, 1974, for its successful use in semantic priming research). Hermans et al. (2002) and Wentura (1998, 2000) only found evidence of evaluative priming in the lexical decision (word/non-word) task. However, to anticipate, these results can be explained by a further mechanism, *viz*. affective matching (Klauer and Stern, 1992; also see Klauer and Musch, 2002), which we further explain below.

Thus, in short, after initial positive findings, early research was characterized by null findings by positive evidence that can be explained by an alternative theory (i.e., affective matching), and even by reversed findings. This suggested that stimulus valence is not a special categorical construct and that it is only activated if it is response relevant. However, in subsequent years, there were several attempts to save the encoding facilitation hypothesis. This research proceeded from different starting points to identify possible moderator variables.

One approach of testing the encoding facilitation hypothesis with more scrutiny started from the premise that standard tasks might only involve superficial processing of the prime concepts, but that deeper processing may be needed to reveal encoding facilitation. For example, Spruyt et al. (2002) argued that early evaluative priming studies with the pronunciation task failed because word stimuli are only processed by a lexical executive system, bypassing semantic analysis (as argued by Glaser and Glaser, 1989), whereas pictures have privileged access to the semantic system (Glaser, 1992). In three experiments, the authors systematically varied the type (i.e., picture vs. word) of primes and targets. They found replicable evaluative priming of naming responses when pictures were used as primes, but not when words were used as primes. Significant effects with picture primes were also found by Duckworth et al. (2002, Exp. 1) and Giner-Sorolla et al. (1999, Exp. 2). Spruyt and Hermans (2008) replicated the picture-based effect again; interestingly, they found the effect already (and only) in the first block of trials, that is, on first occurrence of each stimulus. This is not unimportant to mention, because, in semantic priming research, non-repetition of materials is seen as a precondition of attributing priming effects unequivocally to semantic long-term memory retrieval (see, e.g., McNamara, 2005). Moreover, the result could also be interpreted in terms of Itkes and Kron (2019) distinction between semantic and affective valence. In a study with the affective Simon paradigm (Itkes et al., 2017), they showed that effects based on affective valence fade with habituation, whereas effects based on semantic valence did not. Thus, pictorial scenes might trigger an affective response (as generally observed for such pictures; Bradley et al., 2001), and this affective response facilitates access to congruent concepts. However, Klauer et al. (2016) pointed out that the comparison of evaluatively congruent vs. incongruent pairings in the study by Spruyt and Hermans was confounded with semantic relationships (i.e., on average, congruent prime-target pairs were semantically more related beyond valence congruency than incongruent pairs). In a preregistered replication study that avoided this confound, Klauer et al. found no significant evaluative priming effect. De Houwer et al. (2001) used word-word pairs in a pronunciation task with degraded (e.g., % U% G% L% Y%) and non-degraded targets. They argued that degraded presentation increased semantic processing of the word; hence, a potential facilitation by evaluative congruency might be more impactful. Indeed, an evaluative priming effect was found for degraded targets only. In a similar vein, De Houwer and Randell (2004) conducted two experiments with a conditional pronunciation task. That is, experimental trials were embedded in a sequence of filler trials with targets of a different category than the experimental targets (i.e., non-words, Exp. 1; occupation names, Exp. 2). The participants were instructed to pronounce the targets except when stimuli were non-words or occupations, respectively, that is, they were forced to semantically process the targets. The authors found significant evaluative priming effects in both experiments; however, they acknowledged that their effects may not reflect encoding facilitation, as the participants may have relied on a backward checking mechanism (Neely et al., 1989): Detecting affective congruency between prime and target was a valid indication of trial type (i.e., predictive of the decision "yes, target has to be pronounced") since all filler targets were neutral. Thus, the obtained congruency effect might reflect a faster decision to pronounce the target in case of congruent pairs rather than indicate greater accessibility of the target due to the overlap in valence. (De Houwer and Randell, 2002, Exp. 2) found an evaluative priming effect (with word-word pairs) in a pronunciation task if the participants were instructed "to pay attention to the first word." Thus, taken together, pronunciation tasks using pictorial primes or enforcing deeper encoding of target and/or prime yielded some replicable evidence of evaluative priming. Because of the many unclear and null results, however, this research has not provided a comprehensive explanation of S-S-based priming effects.

Dimensional Attention and Goal Relevance

A more fundamental hypothesis about the validity and constraints of the encoding facilitation hypothesis has been proposed by Spruyt et al. in the last decade. The basic assumption is that "the semantic analysis of task-irrelevant stimuli is modulated by feature-specific attention allocation" (Spruyt et al., 2009, p. 37); that is, evaluative features are assumed to be processed only if attention is directed toward the evaluative dimension (see Kiefer and Martens, 2010, for a comparable assumption in semantic priming research). In three experiments, Spruyt et al. (2009) used a task-switching design to test their assumption. We will focus on Experiment 3 (which resolved some problems with the first two experiments): Depending on a cue (i.e., presence or absence of a green rectangle around the target), the participants had to pronounce target words or categorize them. Categorization trials constituted the majority (75%) of trials. In one group, the participants had to evaluate the targets in the categorization trials (i.e., label them as "positive" or "negative"); thus, the attentional focus was on evaluative features, even–according to the authors–in pronunciation trials. In the second group, the participants had to categorize the targets as objects or people; thus, the attentional focus was on semantic animacy features, even in pronunciation trials. Importantly, prime-targets pairs varied according to evaluative and semantic (i.e., object vs. person) congruency. The pattern of priming effects was in line with hypotheses: In the evaluative focus condition, pronunciation trials produced an evaluative priming effect but no semantic priming effect; the pattern was reversed for the semantic focus condition. Spruyt et al. (2012) replicated this evaluative priming effect even with masked primes, whereas Becker et al. (2016) reported two replication failures. A meta-analysis across all five experiments yielded a significant evaluative priming effect of $d = 0.39$.

Spruyt et al. (2007, Exp. 1) used the same basic design as Spruyt et al. (2012) but with an animal/object categorization task instead of pronunciation and pictures as both primes and targets. Again, a significant evaluative priming effect emerged in the non-evaluative trials (i.e., animal/object categorization trials) if the majority of trials were evaluative (i.e., valence-based categorization trials); the effect was not found if only a minority of trials was evaluative. Experiment 2 used words as primes and targets, and all trials required non-evaluative categorization (living vs. non-living); in addition, the participants were instructed to count the number of positive and negative primes to direct attention to the evaluative dimension. Again, a significant evaluative priming effect was found. In the same vein, Everaert et al. (2011) found an evaluative priming effect (in experimental trials) with pictures as primes and words as targets if filler trials (which constituted the majority of trials) used valent stimuli as well, but not if filler trials used neutral stimuli. An independent conceptual replication of Spruyt et al. (2007, Exp. 1) was published by Kawakami and Miura (2019), who used a mouse-tracking procedure (i.e., the participants categorized items by moving the mouse cursor to a response button). The evaluative priming effect found in non-evaluative trials was impressively large. However, a general critique of these task-switching paradigms (with a majority/minority of evaluation trials) is that the task set of the majority trials might carry over to the minority trials, thereby producing congruency effects. Consequently, the goal-dependent nature of evaluative priming effects, which requires response relevance (at least in part of the task), still holds for these tasks.

Therefore, Werner and Rothermund (2013) introduced a further paradigm that ensures attention is focused on the evaluative dimension without requiring evaluative decisions (or task switching). This paradigm includes trials with evaluatively neutral targets in addition to standard trials (with the variation of prime and target valence). Participants are asked to categorize targets as either valent (i.e., positive or negative) or neutral; thus, attention is on the evaluative dimension without valence discrimination (i.e., positive vs. negative) itself being response relevant. Importantly, the crucial congruency of prime and target valence is irrelevant for the response, which is identical for all valent targets; thus, explanations of an evaluative priming effect in terms of response facilitation or interference are ruled out. While the two experiments in Werner and Rothermund's study yielded null results, their research inspired a remarkable debate about the intricacies of such experiments (Rothermund and Werner, 2014; Spruyt, 2014), but it would be going too far to describe it here. Spruyt and Tibboel (2015) pointed out that some parameters (e.g., the inter-trial interval) used by Werner and Rothermund (2013) were non-standard; replicating their experiment with standard parameters yielded a significant evaluative congruency effect. However, Werner et al. (2018) emphasized a potentially important difference between their original experiment and the replication by Spruyt and Tibboel (2015): Whereas, the former authors used a fixed assignment of primes to targets to avoid any semantic or associative relationships (beyond valence congruency), the latter authors opted for random assignment. This solution entails the risk of a confound of evaluative congruency with semantic/associative relationship. Controlling for semantic associations, Werner et al. did not find an evaluative congruency effect.

In line with this emerging view that the goal relevance of the evaluative dimension is important was a finding by Gast et al. (2014), who introduced a new paradigm to test the encoding facilitation hypothesis. The authors proposed that resolving an evaluative mismatch of prime and target consumes cognitive processing resources, such that a mismatch distracts from *any* task that participants concurrently engage with (see Hermans et al., 1998, for an earlier suggestion of this idea). Gast et al. presented a sequence of two valent images, and a letter (X or Y) that was superimposed on the latter image had to be categorized. In two experiments, letter categorization was slower if pictures were affectively incongruent, but only if a secondary task set an evaluative context.

Spruyt et al. (2018) added a further twist to the discussion: They used a pronunciation task with word targets and picture primes. Pronunciation trials were mixed with a minority of so-called goal-induction trials that were identifiable by the letters

"E" or "F" presented, instead of a target word. The participants were instructed and trained to categorize the letter in these trials *via* keypress, but only if one of four specific pictures preceded the letter. Thus, a small set of two positive and two negative pictures – randomly selected for each participant – acquired "goal relevance." Remarkably, an evaluative priming effect was only found for trials with goal-relevant pictures as primes. Moreover, in line with Spruyt and Hermans (2008), the effect was restricted to the first block.

Taken together, these studies suggest that attention and some level of response relevance of the evaluative dimension are prerequisites for evaluative priming effects to emerge. Without response relevance, positive findings are more likely to be based on semantic/associative relatedness or – in the case of the Gast et al. study – are potentially based on a conflict reduction process (see also below). Thus, stimulus valence does not seem to cause activation spread to all concepts of the same valence. An exception might be affective pictures in studies that only yielded effects in the first blocks; such effects might be based on an affective response (Itkes et al., 2017). However, there has not been a systematic investigation of this assumption in the evaluative priming literature. A remarkable new line of research might be constituted by the study of Spruyt et al. (2018), who found evaluative priming effects only for those stimuli that had general goal relevance. Implications for models of the mental representation of stimulus valence are discussed further below.

Mutual Facilitation

A different perspective on the underlying activated representations was taken by Wentura et al. (Wentura and Rothermund, 2003; Wentura and Frings, 2008; Schmitz and Wentura, 2012; Schmitz et al., 2014). They assumed that encoding facilitation might be a bidirectional process. Specifically, a prime might facilitate processing of a congruent target; in addition, however, a target might facilitate access to a congruent prime as well. Thus, it is assumed that both representations are activated in parallel. Testing this idea becomes interesting when prime and target are associated with the same valence, but different response tendencies. In this case, we might expect *reversed* priming effects (e.g., see Glaser and Banaji, 1999) because two concepts that mutually facilitate each other (because of the shared valence) compete for response determination. Wentura and Frings (2008) found evidence for this by using a picture/picture version of the priming paradigm and a naming task. In this version of the paradigm, target naming must be trained before the priming phase to achieve fast and unequivocal naming. Wentura and Frings added the primes to this training phase as well, with half of them trained with the adequate object name and half with a random response (i.e., consecutive numbers). Thus, primes and targets were presented before the priming phase, and object labeling was trained (i.e., the correct response was learned) or untrained for the primes. Interestingly, the untrained primes caused a positive priming effect, whereas the effect numerically reversed for the trained, response-bound primes. By changing the stimulus onset asynchrony to −80 ms (i.e., the prime was presented on top of the bigger target picture 80 ms after the target onset), Schmitz and Wentura (2012) found a symmetric pattern of a positive priming effect for untrained primes and a significant reversed effect for trained primes, indeed indicating that prime and target might be activated in parallel. This parallel activation is difficult to explain with encoding facilitation, which assumes that activation spreads from one concept to others (such that activation dissipates for the prime if the target becomes activated). Schmitz and Wentura corroborated this pattern of results with a semantic classification task (i.e., persons vs. animals) using pictures (Exp. 2) or words (Exp. 3; also see Schmitz et al., 2014). Interestingly, the authors found that the to-be-expected *semantic* congruency effect was moderated by *evaluative* congruency, which implies that stimulus valence was processed despite being response irrelevant. These effects can only be explained with mutual facilitation of prime and target.

In sum, the available evidence paints a rather complex picture. Evaluative priming effects with the evaluative categorization task are a robust phenomenon, which can, however, be explained with response compatibility as the underlying mechanism. This makes it difficult to gain much insight into the underling mental representation of valence and the links between concepts. Evidence accrued with non-evaluative categorization, pronunciation, and lexical decision tasks sheds more light on the underlying representations – however, the results are complicated: With some caution, we can state that there seem to be conditions that yield evaluative congruency effects with these tasks. Specifically, attention or goal relevance and sufficiently deep (i.e., semantic) processing of the evaluative dimension seem to be required for such effects to emerge. Moreover, at least in two cases, evaluative congruency effects were found to depend on additional semantic associations between prime and target (Klauer et al., 2016, in a replication of Spruyt and Hermans, 2008; Werner et al., 2018, in a replication of Spruyt and Tibboel, 2015). Thus, associative relationships at least boost the evaluative priming effects. Given this overall, how can the mental representation of valence be modeled?

Theoretical Models of S-S-Based Evaluative Priming

For a long time, the guiding hypothesis of S-S-based evaluative priming was encoding facilitation of the target concept by an evaluatively congruent prime *in long-term semantic memory*.

This encoding facilitation was first explained with the theory of semantic networks (Anderson, 1983; Collins and Loftus, 1975; for a recent discussion, see Kumar et al., 2021): Semantic concepts are assumed to be represented by nodes in a network; links between nodes symbolize semantic or associative relationship. Processing a word leads to the "activation" of the corresponding node and a process of "spreading activation" that increases the activation level of related concepts *via* the links. Hence, these concepts become more accessible for processing compared to a baseline condition. The semantic network theory was adapted by Bower (1991) to explain mood congruity effects and was also embraced by evaluative priming researchers. The guiding assumption is that a "positivity node" and a "negativity node" exist, linking all positively and negatively connoted concepts, respectively. However, it can be argued that a semantic network will be functional only if the amount of spreading activation is constrained by a reasonable number of links emanating from the source node – otherwise, activation (e.g., from a generic positivity or negativity node) of target concepts will become too small to be measurable (a "fan effect"; see Anderson, 1974; also see Hermans et al., 1996; Wentura, 1999; Spruyt et al., 2002). There are several additional problems: First, semantic priming experiments use unique primes and targets to unequivocally interpret observed effects in terms of semantic long-term memory retrieval (see, e.g., McNamara, 2005); by contrast, most evaluative priming experiments use comparably small prime and target sets with a high repetition rate (for exceptions, see Klauer and Musch (2001), Exp.1; Spruyt and Hermans (2008). Second, when comparing evaluative and semantic priming effects, one should keep in mind that evaluative priming can be considered a special case of semantic priming, that is, priming of category members by category coordinates that are not (necessarily) associated (e.g., *robin-eagle* as exemplars for the category *bird*). For semantic priming, Hutchison (2003, p. 804) summarized that "the overall evidence of [semantic] priming from category coordinates is weak" (but see Lucas, 2000). If this conclusion holds, the *a priori* expectation of finding S-S-based evaluative priming effects is rather low. Or in other words: If there is some evidence for S-S-based evaluative priming effects, we should carefully discuss whether such effects should be interpreted as semantic priming effects because, in general, evidence of semantic priming from category coordinates is weak. Semantic network theory should thus be seen more as a *metaphor* that illustrates the phenomenon of priming effects, but it does not really explain the phenomenon in a deeper sense.

More promising (at least at first glance) are parallel-distributed models of priming (e.g., Masson, 1995; McRae et al., 1997). In these models, each semantic concept corresponds to a specific pattern of activated processing units, and the semantic relatedness between concepts is determined by the number of shared activated units. This conceptualization is more far-reaching than that of semantic networks, since parallel-distributed models contain assumptions about the emergence of the structures (*via* learning mechanisms). Priming effects are explained by a faster transition from the pattern representing the prime concept to a semantically related one (i.e., the pattern representing the related target) than to an unrelated pattern because the shared units are already in the appropriate mode of activation. Although this model was developed to explain semantic priming (see McNamara, 2005), it is prima facie perfectly suited to account for S-S-based evaluative priming as well, under the assumption that a considerable part of the activation pattern represents the evaluative features of a concept (see Wentura, 1999, 2000; Spruyt et al., 2002).

However, there are two main problems associated with this hypothesis: First, Spruyt et al. (2007) pointed out that, according to simulation studies by Masson (1995), the overlap of prime and target representations must be very large to account for priming effects; this would mean that the evaluative part of the stimulus representation must be an implausibly large part of the representation of concepts. The second problem of parallel-distributed models of priming was discussed by Schmitz and Wentura (2012). They argued that some findings in evaluative priming research can only be explained by parallel activation of prime and target representations (e.g., reversed priming effects with the pronunciation task; Glaser and Banaji, 1999; Wentura and Frings, 2008; Schmitz and Wentura, 2012). As parallel-distributed models assume that there is a transition from the prime pattern to the target pattern (which is faster if concepts are related), the prime representation should no longer be accessible after this transition. Scherer and Wentura (2018) directly tested the assumption of parallel activation using a post-cue priming task (i.e., a cue following a brief and masked presentation of two words indicates which of the two words has to be identified). Results were in line with the assumption of parallel activation, suggesting that parallel-distributed models might also not be suitable to explain the activations underlying evaluative priming effects.

Thus, in a nutshell, S-S-based evaluative priming effects are not (easily) explainable with the two well-known long-term memory models that were used to explain semantic priming effects. So how can we explain them? In our view, three mechanisms need to be considered to better understand the effects. In our view, these mechanisms (and the activated prime and target representations) have to be modeled as operating in working memory, not in long-term semantic memory.

We already mentioned *affective matching theory* by Klauer and Stern (1992): According to this theory, if two evaluative stimuli are presented, the two valences are (a) involuntarily processed and (b) involuntarily compared; this comparison results in an affective match or mismatch. Depending on the context, match or mismatch influences ongoing behaviors, primarily *via* affirmative or negative response tendencies. Since binary decision tasks often involve affirmation (e.g., "yes, this is a word"), congruency effects (i.e., faster responses in the match case) are predicted. However, the touchstone of such a theory is the case of negation. If a task set can be changed to a negation situation *via* instructions (e.g., in a lexical decision: "no, this is *not* a senseless letter string"), the effect should reverse. This was, indeed, observed in some studies (see Klauer and Stern, 1992; Wentura, 1998, 2000; Klauer and Musch, 2002; Rothermund and Werner, 2014). Affective matching mechanisms are especially likely to govern the effects if *pairs* of stimuli have to be compared on a feature that is orthogonal to valence [e.g., do the words share lettercase (lower or uppercase)?]. Klauer and Musch (2002) tested this in four experiments with different aspects, and found results that were in line with matching theory: If the "same" response was (by instruction) affirmative, it was facilitated by the task-irrelevant evaluative congruency of the stimuli (compared to incongruency); if the "different" response was affirmative, the effect was eliminated and tended to be reversed. Note that the matching mechanism is about the comparison of two activated concepts; hence, an obvious suggestion is to locate it in working memory. However, effects found with the pronunciation task are problematic for this theory because pronouncing a word is not a judgment of affirmation or negation.

Such effects might be explained by a further mechanism, that is, *affective motivational conflict reduction*. Specifically, Hermans et al. (1998) suggested that the affective incongruency of two stimuli *per se* slows down system behavior due to an inherent processing conflict. They argued that automatic evaluation results in action tendencies of approach (for positive stimuli) or avoidance (for negative stimuli). In case of a mismatch, the action tendencies are in conflict; this conflict blocks behavior until it is resolved. However, it might be sufficient to assume that an affective match or a mismatch serves as a signal: A match might signal: "everything as expected; go on!", whereas a mismatch might signal: "violation of expectation; clarify!" However, in four experiments with a color-naming task with evaluatively congruent or incongruent prime-target pairs, Hermans et al. did not find evidence for their claim (see also Rothermund and Wentura, 1998). The evidence presented by Gast et al. (2014; see above) on the other hand fits the affective-motivational account. This theory is, in principle, able to explain congruency effects in all evaluative priming tasks (although it does struggle to explain reversed effects). Its strength, however, lies in the potential to explain effects if *both* valent stimuli are *task irrelevant* (in contrast to affective matching theory, which is best suited to explain effects if both stimuli are task relevant). Again, the conflict reduction account is about the comparison of two activated concepts; hence, again, an obvious suggestion is to locate it in working memory.

A third candidate account relates to processes based on *mutual facilitation or inhibition in working memory*. Wentura et al. (Wentura and Rothermund, 2003; Wentura and Frings, 2008; Schmitz and Wentura, 2012; Schmitz et al., 2014) emphasized the parallel activation of prime and target concepts and their mutual facilitation and/or inhibition (see Section Mutual facilitation above). Depending on task context, parallel activation can have different consequences (Wentura and Rothermund, 2003). Task context can be characterized by two dichotomies: The first dichotomy refers to whether participants have to categorize the valence (in an evaluation task, or, more generally, a response priming design) or not (in tasks such as pronunciation or a lexical decision, or, more generally, a semantic priming design). The second dichotomy refers to whether or not the prime serves as a distractor competing with the target for determination of the response.

To elaborate on the latter: In the evaluation task, the usual instruction to respond fast might lead to the following processing strategy: "Whenever the stimulus display delivers enough evidence for a positive or negative response, I will press the corresponding key. "One might add: "… whatever the basis for this evaluation might be." That is, participants may not try to separate prime and target (also see the evaluation window account; Klauer et al., 2009). Thus, congruent pairs will facilitate the response because the evidence provided by the stimulus configuration is unambiguous. If, however, the participants follow the strategy to base their decisions on the correct source of evidence only – that is, on target valence – the prime is, in fact, a distractor whose activation hinders correct response formation. Thus, in the case of congruency, the competition between prime and target representations is prolonged by the congruent target facilitating processing of the prime and/or in the case of incongruency, competition is shortened due to the incongruent target inhibiting processing of the prime. If the participants use this strategy, reversed effects (e.g., Wentura and Rothermund, 2003; also see Klauer et al., 2009[4]) – or at least reduced congruency effects (see, e.g., Wentura and Degner, 2010a) – can be found even with the evaluation task.

The pronunciation task can serve as another example: In this task, participants obviously need to focus on the target representation to pronounce the target; thus, in a general sense, the prime is always a distractor. Response execution must, therefore, wait for the activation of the target representation to reach the threshold and activation of the prime representation to be diminished (Houghton and Tipper, 1994). This can be seen as a routine process, given that reading of words almost always occurs in the context of other words. However, in the priming task, we can create, more or less, competition by associating the distractor with automatic response tendencies or not (Wentura and Frings, 2008; Schmitz and Wentura, 2012). If there are no such competing response tendencies, we might find evidence for congruency effects, because a congruent prime helps to establish the activation of the target representation, especially with the usual prime-target onset asynchrony. If, however, the distractor is associated with a competing response tendency, the mutual facilitation in case of congruency can prolong the competition between a distractor and a target, especially with a negative onset asynchrony of prime and target (Schmitz and Wentura, 2012). Interestingly, a working memory model that was developed in a totally independent area of research – the dual-store neuro-computational model of short-term memory (Usher and Cohen, 1999; Haarmann and Usher, 2001; Davelaar et al., 2005, 2006) – proposes mutual facilitation processes based on semantic similarity in working memory. Given this backdrop, Scherer and Wentura (2022) used a change detection task – that is, a typical working memory task – with emotional (i.e., angry and happy) faces as stimuli to investigate the effects of evaluative congruency on working memory performance. They found an (admittedly: weak) effect of evaluative congruency, with better performance on trials with evaluatively congruent compared to incongruent displays. This effect is in line with the assumptions that S-S-based evaluative priming effects might arise from mutual facilitation/inhibition of simultaneously active evaluatively congruent concepts.

In Section Dimensional Attention and Goal Relevance, we emphasized in the concluding paragraph the recent study by Spruyt et al. (2018) who found evaluative priming in the pronunciation task only for those stimuli, which had general goal relevance: Concretely, a small set of two positive and two negative prime pictures indicated that, if a letter instead of a target stimulus will be presented in a given trial, it had to be categorized. (For all other primes, no response to the letter was allowed). We can interpret this finding in the following way: In

[4]Klauer et al. (2009) presented an alternative explanation for reversed effects in the evaluation task. We will not elaborate on this account because our focus in this review is on S-S-based evaluative priming (i.e., evaluative priming in the semantic priming design).

contrast to goal-irrelevant primes, who might produce a very superficial activation of its mental representation, a goal-relevant prime might mandatorily become part of the current working memory to prepare for the potential letter classification task. Given this status, the primes might facilitate congruent targets.

Taken together, the three outlined mechanisms – as different as they might seem at first glance – have much in common: All three approaches are compatible with the assumption that two concepts are activated in parallel and that evaluative congruency vs. incongruency influences further processing while both concepts are activated. It then depends on specific task parameters which of the three mechanisms determines observed effects. That is, we can identify task contexts that produce results that are either more compatible with affective matching, conflict reduction, or mutual facilitation/inhibition. If both stimuli are task relevant, the affective matching process might be dominant; if both stimuli are task irrelevant, the conflict reduction process might be dominant. Finally, if only one stimulus is task relevant, mutual facilitation or inhibition might be the processes that determine evaluative congruency effects. Of course, there is overlap: The affective matching mechanisms might be evoked if only the target is task relevant: However, a clear judgmental context must be given that asks for affirmation or negation; then, affective match or mismatch meddles as well with giving a response (Wentura, 1998, 2000; Klauer and Musch, 2002 Exp. 9). The conflict reduction process might, in principle, be at work in all three task contexts. However, it cannot be the sole account because some results cannot be explained by it (e.g., reversed effects found with the pronunciation task, Schmitz and Wentura, 2012, or the effects found by Klauer and Musch (2002) see above, that fit to affective matching theory).

Finally, to (provisionally) answer the question posed at the outset of this section – "How is valence represented in conceptual representation systems?" – we can conclude that valence connotations are –in a specific sense – a fundamental part of the representation of semantic concepts: Evaluative information will be involuntarily activated in specific contexts, that is, if attention or goal relevance is given. Then, the evaluative aspects of the corresponding concept are accessed (which might be equated with: is activated in working memory). However, the "in a specific sense" constraint should be understood as: There is not much evidence for a conceptual representation system that has valence as a basic organizational feature such that processing of a positively connoted concept renders all other positive concepts more accessible, for example. But if context conditions trigger the activation of evaluative connotations and two concepts are activated in parallel (which might be equated with: both are concurrently represented in working memory), then affective match or mismatch will influence behavior.

CAN FURTHER EMOTION ASPECTS BE INVOLUNTARILY ACTIVATED?

From what has been discussed so far, we might conclude that there is not much special about the processing of stimulus valence in evaluative priming paradigms. Semantic valence seems to be activated, and this happens only under specific processing conditions. Nevertheless, there is evidence suggesting that there could be special aspects in the processing of emotion and affect in evaluative priming. Recently, researchers have started investigating the processing of specific emotions or specific emotion aspects (i.e., possessor/other relevance, potency, specific emotion categories) with the paradigm. From the viewpoint of semantic processing accounts, one might think that the processing of such aspects occurs if they are task relevant and similar to a semantic category. The emerging evidence, however, tells a different story. For example, (Wentura and Degner, 2010a, also see Degner et al., 2007; Degner and Wentura, 2011) examined whether possessor vs. other relevance plays a role in evaluative priming effects. This differentiation was introduced by Peeters (1983) who posited that many trait adjectives are not simply positive or negative, but that the evaluation depends on the perspective of the evaluator. For example, being aggressive is not necessarily bad for the aggressor her/himself but is probably seen as bad by the social environment; that is, it is an other-relevant trait. Likewise, being intelligent is a positive trait for the trait holder, but not necessarily for others, so that this trait is (primarily) possessor relevant. While this distinction is interesting from a social viewpoint, one might assume that some controlled processing or conscious reflection is necessary to make it. On the contrary, however, the studies by Wentura et al. showed that this differentiation is of relevance in evaluative priming paradigms. For example, Wentura and Degner (2010a) observed masked evaluative priming effects in two studies only when prime and target were congruent with regard to *both* valence *and* possessor/other relevance. Of note, their analysis controlled for semantic relatedness, so the differentiation cannot simply be explained by the assumption that concepts of the same relevance type are more closely connected semantically. The authors thus proposed that relevance type is also assessed early and automatically with stimuli for which this differentiation is relevant. Several other studies corroborated this finding, also under masked presentation conditions (Degner et al., 2007; Degner and Wentura, 2011; De Paula Couto and Wentura, 2012).

While the possessor/other relevance differentiation certainly has some appeal, it can possibly be subsumed under a more general point: It might be the case that the specific emotional connotation is activated instead of coarse valence and that only prime-target pairs with matching connotation are processed as congruent (e.g., *incapable* – *depressed* might be perceived as congruent because both relate to sadness). To examine this hypothesis, Rohr et al. (2012) built on work by Carroll and Young (2006), who adapted the evaluative priming paradigm to specific emotions. In this paradigm, participants are given four response options and have to categorize targets according to the specific emotion category (i.e., joy, anger, fear, sadness). Carroll and Young (2006) found priming effects based on emotion-category congruency in four studies using words, facial expressions, and emotional sounds. However, they used an unusually long stimulus-onset asynchrony (SOA; i.e., the time from the prime onset to the target onset) in all but one experiment (i.e., 750 ms; 250 ms in Exp. 4), and analyzed data at a rather coarse level of granularity (i.e., comparing all emotion-congruent to all emotion-incongruent conditions; with the exception of Exp. 4, which excluded the positive emotion). Thus, results could not rule out a certain level of controlled processing, and the degree to which the participants differentiated emotions remained clear: Specifically, did the participants differentiate between all emotion categories? Or did they only differentiate specific emotion aspects, such as the abovementioned possessor/other relevance? Rohr and colleagues (Rohr et al., 2012; Rohr and Wentura, 2014; Wentura et al., 2017) examined the degree of processing differentiation in more detail using a masked version of the emotion priming paradigm. Their analysis focused on the prespecified set of Helmert contrasts: The first contrast treated all positive-positive and negative-negative (e.g., sadness-anger) prime-target combinations as congruent and all combinations of positive and negative valence as incongruent, thereby testing for differentiation of valence. The next contrast compared anger vs. fear/sadness combinations, excluding all trials containing the only positive emotion, joy. Specifically, anger/anger trials, as well as all combinations of fear/sadness, were treated as congruent, and all trials involving anger *and* fear or sadness as incongruent, with the rationale, that the two latter emotions both convey the same type of relevance (i.e., possessor relevance; alternatively, this differentiation can be interpreted as an assessment of potency, Neumann et al., 2020; or coping ability, Rohr et al., 2020). The last contrast compared the remaining two emotions (i.e., all trials involving fear/sadness only). Rohr et al. (2012) found emotion priming effects beyond valence, concretely a differentiation of anger vs. fear/sadness, but no differentiation of each specific emotion category (i.e., fear and sadness were not differentiated). The authors interpreted the finding as evidence for the extraction of valence and possessor/other relevance. What distinction exactly underlies this intermediate differentiation is a matter for debate; however, in any case, the finding is very interesting for the issue of underlying processes: From a semantic processing (and response compatibility) viewpoint, emotion-specific priming effects should have emerged. Thus, the observed intermediate differentiation cannot easily be explained by semantic processing accounts. By contrast, the findings provide preliminary evidence that early processing can extract more than just valence, for social and affect-related reasons (also see Rohr et al., 2020). Of note in this regard, different patterns of early differentiation beyond valence have been observed. In other studies (Rohr and Wentura, 2014; Rohr et al., 2015), an intermediate differentiation of anger/fear vs. sadness emerged, which can be interpreted as a differentiation along the arousal or threat dimension, albeit with a different paradigm (Rohr et al., 2015) or using specific stimuli (i.e., low spatial frequency-filtered facial expressions; Rohr and Wentura, 2014). This inconsistency leaves room for an interpretation along methodological lines: Possibly, individuals can only focus on two stimulus dimensions under fast, unintentional processing conditions such that only two response categories yield specific effects (please note that clearly visible primes yielded congruency effects for all four emotion categories in; Rohr and Wentura, 2014; Rohr et al., 2015).

Wentura and Rohr (2018) tackled this issue with a further version of evaluative priming – the "leave-one-out" paradigm. In four studies, the participants completed emotion priming tasks with only two target categories (e.g., anger vs. sadness); however, there were three categories of negative masked primes (i.e., anger, fear, sadness), with primes drawn from a different set than targets. This setup allows examination of the impact of a masked prime that is not part of the task set, for which, presumably, no response trigger is built up. In a nutshell, results supported an interpretation in terms of emotion-specific processing, however, with task-relevance playing a primary role: The response-relevant primes always yielded clear congruency effects, whereas priming effects were approximately halved with the "left-out" prime. This suggests that the left-out prime category is differentiated from the other two response-relevant categories, but that it does not lead to the activation of a "response trigger," given that no response category is associated (see Kunde et al., 2003, for the action trigger account). These results mesh well with the results of Neumann and Lozo (2012), who observed emotion-specific priming effects for masked disgust and fear stimuli in three studies with a disgust vs. fear categorization task.

Recently, Neumann et al. (2020) have also adapted the evaluative priming paradigm to study the automatic processing of potency, which the authors defined as an umbrella term for power, dominance, control, and related physical concepts, such as weight and size. After valence and arousal, potency is a further dimension defining the affective semantic space (Osgood et al., 1957), but only few studies have targeted this dimension explicitly. In three studies, Neumann and colleagues showed that a potency categorization task (i.e., classify the target as strong vs. weak) reliably elicits potency priming effects, but only under specific conditions (i.e., with a short SOA, a response window procedure, with adjectives but not nouns). They also showed that the effects depend on task instructions: With a potency categorization task, no effects for valence (Exp. 1a; Exp. 3) or arousal (Exp. 1b) were observed even though the stimuli varied on these dimensions. In a valence categorization task (Exp. 3), the potency priming effect was reduced but still significant. The authors attributed this latter finding to the salience of this dimension in their stimulus pool. Importantly, the potency priming effects still held when controlling for semantic overlap, suggesting that the processing of this dimension is "an important mechanism in addition and beyond valence that regulates social behavior and prepares the organism to act adaptively" (Neumann et al., 2020, p.2).

Thus, taken together, studying emotion-related aspects or dimensions with priming paradigms shows that more than valence can be extracted early, and even under masked presentation conditions. Similar to the evaluative effects discussed earlier, the effects seem to depend on specific task parameters. However, the effects cannot be explained solely by semantic associations; whether certain aspects or dimensions are processed seems to be related to their social importance. Perhaps, the most interesting finding in this regard is that processing of specific emotions under short and masked presentation conditions seems to be confined to two dimensions only, even if the task requires processing of the specific emotion. Given the results of Wentura and Rohr (2018), this two-dimensional pattern might be related to task complexity (i.e., four-response options). Alternatively, it might reflect a more general restriction of processing: emotion research and social psychological research have often found two important dimensions (Russell and Barrett, 1999; Fiske et al., 2002; Oosterhof and Todorov, 2008), and recent research has also suggested that current goals moderate which dimensions or aspects are attended to (Nicolas et al., 2022). In this respect, the results might also align with the abovementioned working memory account of evaluative priming. Perhaps, only two dimensions are kept active in working memory under implicit processing conditions. However, more research is needed to improve our understanding of these matters.

SUMMARY AND CONCLUSION

To summarize, research using evaluative priming paradigms has revealed that information in our environment is evaluated automatically; that is, prime valence is processed unintentionally, efficiently, and also non-consciously, if (and only if) the evaluative dimension is relevant to the task goal. Moreover, the underlying mental representations clearly seem to be semantic in nature. "Hot" processing does not seem to be involved. A few studies, however, give rise to the assumption that the affective state can influence priming effects (Storbeck and Clore, 2008; Foroni and Semin, 2012; Lemonnier and Alexopoulos, 2019). Furthermore, some studies with picture primes, which found effects in the first block only, leave open whether an affective response might contribute to evaluative priming effects in studies employing pictures; however, systematic research into this question is still lacking. Thus, despite three decades of research with the paradigm, the complex pattern of evidence still does not allow an authoritative answer to the question of how to model the mental representations and activated valence of prime and target. The proposed long-term memory models all struggle to explain the evidence in its entirety. We proposed a working memory account, which, however, requires further development.

Does this mean that there is nothing special about the processing of stimulus valence and that it should be equated to other semantic categories? We do not think so, even if the activated valence reflects factual knowledge (i.e., semantic valence). As described in Section Can Further Emotion Aspects be Involuntarily Activated?, observed effects seem related to the social or personal relevance of the information, and suggest that there is some early extraction of affect-related information beyond valence. It could be that this automatic processing is restricted to the processing of two socially relevant dimensions only, thereby converging with other social theories (e.g., Nicolas et al., 2022), but, again, further research is necessary to corroborate this notion.

Thus, to conclude, the evaluative priming paradigm has revealed much about the link between emotion and cognition: Semantic affective knowledge is automatically activated when encountering stimuli in a context in which the evaluative dimension is relevant. It is even possible that additional socially relevant emotion aspects are activated in this situation. Existing research suggests that the effects are not explainable with theories of semantic priming. In that sense, the processing of stimulus valence might still be special. How *exactly* it is special, however, has not been specified in a comprehensive theoretical account just yet.

AUTHOR CONTRIBUTIONS

MR and DW conceived the review and wrote substantial parts of the manuscript text. Both authors contributed to the article and approved the submitted version.

FUNDING

We acknowledge support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) and Saarland University within the funding programme Open Access Publishing.

REFERENCES

Anderson, J. R. (1974). Retrieval of propositional information from long-term memory. *Cognitive psychology* 6, 451–474. doi: 10.1016/0010-0285(74)90021-8

Anderson, J. R. (1983). A spreading activation theory of memory. *J. Verbal. Learn Verbal. Behav.* 22, 261–295. doi: 10.1016/S0022-5371(83)90201-3

Bargh, J. A., Chaiken, S., Raymond, P., and Hymes, C. (1996). The automatic evaluation effect: unconditional automatic attitude activation with a pronunciation task. *J. Exp. Soc. Psychol.* 32, 104–128. doi: 10.1006/jesp.1996.0005

Barrett, L. F. (2014). The conceptual act theory: a précis. *Emot. Rev.* 6, 292–297. doi: 10.1177/1754073914534479

Becker, M., Klauer, K. C., and Spruyt, A. (2016). Is attention enough? A re-examination of the impact of feature-specific attention allocation on semantic priming effects in the pronunciation task. *Atten. Percept. Psychophys.* 78, 396–402. doi: 10.3758/s13414-015-1046-7

Berner, M. P., and Maier, M. A. (2004). The direction of affective priming as a function of trait anxiety when naming target words with regular and irregular pronunciation. *Exp. Psychol.* 51, 180–190. doi: 10.1027/1618-3169.51.3.180

Boster, J. S. (2005). "Emotion categories across languages," in *Handbook of Categorization in Cognitive Science*, 2nd Edn, eds H. Cohen and C. Lefebvre (Saint Louis, MO: Elsevier Science), 313–352.

Bower, G. H. (1991). "Mood congruity of social judgments," in *Emotion and Social Judgments*, eds J. P. Forgas (Oxford: Pergamon Press), 31–53.

Bradley, M. M., Codispoti, M., Cuthbert, B. N., and Lang, P. J. (2001). Emotion and motivation I: defensive and appetitive reactions in picture processing. *Emotion.* 1, 276–298.

Carroll, N. C., and Young, A. W. (2006). Priming of emotion recognition. *Q. J. Exp. Psychol.* 58, 1173–1197. doi: 10.1080/02724980443000539

Collins A. M., and Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychol. Rev.* 82, 407–428. doi: 10.1037/0033-295X.82.6.407

Cunningham, W. A., Zelazo, P. D., Packer, D. J., and Van Bavel, J. J. (2007). The iterative reprocessing model: a multilevel framework for attitudes and evaluation. *Soc. Cogn.* 25, 736–760. doi: 10.1521/soco.2007.25.5.736

Davelaar, E. J., Goshen-Gottstein, Y., Ashkenazi, A., Haarmann, H. J., and Usher, M. (2005). The demise of short-term memory revisited: Empirical and computational investigations of recency effects. *Psychol. Rev.* 112, 3–42. doi: 10.1037/0033-295X.112.1.3

Davelaar, E. J., Haarmann, H. J., Goshen-Gottstein, Y., and Usher, M. (2006). Semantic similarity dissociates short-from long-term recency effects: Testing a neurocomputational model of list memory. *Mem. Cognit*. 34, 323–334. doi: 10.3758/BF03193410

De Houwer, J. (2003). On the role of stimulus-response and stimulus-stimulus compatibility in the stroop effect. *Mem. Cognit*. 31, 353–359. doi: 10.3758/BF03194393

De Houwer, J., and Hermans, D. (1999). *Nine Attempts of Find Affective Priming of Pronunciation Responses: Effects of SOA, Degradation, and Language*. Paper presented at the Tagung der Fachgruppe Sozialpsychologie, Kassel.

De Houwer, J., Hermans, D., and Eelen, P. (1998). Affective and identity priming with episodically associated stimuli. *Cogn. Emot*. 12, 145–169. doi: 10.1080/026999398379691

De Houwer, J., Hermans, D., Rothermund, K., and Wentura, D. (2002). Affective priming of semantic categorisation responses. *Cogn. Emot*. 16, 643–666. doi: 10.1080/02699930143000419

De Houwer, J., Hermans, D., and Spruyt, A. (2001). Affective priming of pronunciation responses: effects of target degradation. *J. Exp. Soc. Psychol*. 37, 85–91. doi: 10.1006/jesp.2000.1437

De Houwer, J., and Randell, T. (2002). Attention to primes modulates affective priming of pronunciation responses. *Exp. Psychol*. 49, 163–170. doi: 10.1026//1618-3169.49.3.163

De Houwer, J., and Randell, T. (2004). Robust affective priming effects in a conditional pronunciation task: evidence for the semantic representation of evaluative information. *Cogn. Emot*. 18, 251–264. doi: 10.1080/02699930341000022

De Houwer, J., and Tucker Smith, C. (2013). Go with your gut!: Effects in the affect misattribution procedure become stronger when participants are encouraged to rely on their gut feelings. *Soc. Psychol.* 44, 299–302. doi: 10.1027/1864-9335/a000115

De Paula Couto, M. C. P., and Wentura, D. (2012). Automatically activated facets of ageism: Masked evaluative priming allows for a differentiation of age-related prejudice. *Eur. J. Soc. Psychol*. 42, 852–863. doi: 10.1002/ejsp.1912

De Saedeleer, L., and Pourtois, G. (2016). Evaluative priming reveals dissociable effects of cognitive versus physiological anxiety on action monitoring. *Emotion* 16, 498. doi: 10.1037/emo0000149

Degner, J., and Wentura, D. (2011). Types of automatically activated prejudice: assessing possessor-versus other-relevant valence in the evaluative priming task. *Soc. Cogn*. 29, 182–209. doi: 10.1521/soco.2011.29.2.182

Degner, J., Wentura, D., Gniewosz, B., and Noack, P. (2007). Hostility-related prejudice against Turks in adolescents: masked affective priming allows for a differentiation of automatic prejudice. *Basic Appl. Soc. Psych*. 29, 245–256. doi: 10.1080/01973530701503150

Dehaene, S., Naccache, L., Le Clec,'H., G., Koechlin, E., Mueller, M., et al. (1998). Imaging unconscious semantic priming. *Nature* 395, 597–600. doi: 10.1038/26967

Duckworth, K. L., Bargh, J. A., Garcia, M., and Chaiken, S. (2002). The automatic evaluation of novel stimuli. *Psychol. Sci.* 13, 513–519. doi: 10.1111/1467-9280.00490

Dukes, D., Abrams, K., Adolphs, R., Ahmed, M. E., Beatty, A., Berridge, K. C., et al. (2021). The rise of affectivism. *Nat. Hum. Behav*. 5, 816–820. doi: 10.1038/s41562-021-01130-8

Everaert, T., Spruyt, A., and De Houwer, J. (2011). On the (un) conditionality of automatic attitude activation: the valence proportion effect. *Can. J. Exp. Psychol*. 65, 125. doi: 10.1037/a0022316

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., and Kardes, F. R. (1986). On the automatic activation of attitudes. *J. Pers. Soc. Psychol.* 50, 229–238. doi: 10.1037/0022-3514.50.2.229

Fiske, S. T., Cuddy, A. J., Glick, P., and Xu, J. (2002). A model of (often mixed) stereotype content: competence and warmth respectively follow from perceived status and competition. *J. Pers. Soc. Psychol.* 82, 878–902. doi: 10.1037/0022-3514.82.6.878

Foroni, F., and Semin, G. R. (2012). Not all implicit measures of attitudes are created equal: evidence from an embodiment perspective. *J. Exp. Soc. Psychol.* 48, 424–427. doi: 10.1016/j.jesp.2011.08.015

Gast, A., Werner, B., Heitmann, C., Spruyt, A., and Rothermund, K. (2014). Evaluative stimulus (in) congruency impacts performance in an unrelated task. *Exp. Psychol*. 61, 187–95. doi: 10.1027/1618-3169/a000238

Gawronski, B., Cunningham, W. A., LeBel, E. P., and Deutsch, R. (2010). Attentional influences on affective priming: Does categorisation influence spontaneous evaluations of multiply categorisable objects? *Cognit. Emot.* 24, 1008–1025. doi: 10.1080/02699930903112712

Giner-Sorolla, R., Garcia, M. T., and Bargh, J. A. (1999). The automatic evaluation of pictures. *Soc. Cogn*. 17, 76–96. doi: 10.1521/soco.1999.17.1.76

Glaser, J., and Banaji, M. R. (1999). When fair is foul and foul is fair: reverse priming in automatic evaluation. *J. Pers. Soc. Psychol*. 77, 669. doi: 10.1037/0022-3514.77.4.669

Glaser, W. R. (1992). Picture naming. *Cognition* 42, 61–105. doi: 10.1016/0010-0277(92)90040-O

Glaser, W. R., and Glaser, M. O. (1989). Context effects in stroop-like word and picture processing. *J. Exp. Psychol. Gen*. 118, 13. doi: 10.1037/0096-3445.118.1.13

Haarmann, H., and Usher, M. (2001). Maintenance of semantic information in capacity-limited item short-term memory. *Psychon. Bull. Rev*. 8, 568–578. doi: 10.3758/BF03196193

Harnad, S. (2005). "To cognize is to categorize," in *Handbook of Categorization in Cognitive Science*, 2nd Edn, eds H. Cohen and C. Lefebvre (Saint Louis, MO: Elsevier Science), 21–54.

Hermans, D., De Houwer, J., and Eelen, P. (1994). The affective priming effect: Automatic activation of evaluative information in memory. *Cogn. Emot*. 8, 515–533. doi: 10.1080/02699939408408957

Hermans, D., De Houwer, J., and Eelen, P. (1996). Evaluative decision latencies mediated by induced affective states. *Behav. Res. Ther*. 34, 483–488. doi: 10.1016/0005-7967(96)00017-4

Hermans, D., De Houwer, J., and Eelen, P. (2001). A time course analysis of the affective priming effect. *Cogn. Emot.* 15, 143–165. doi: 10.1080/02699930125768

Hermans, D., Smeesters, D., De Houwer, J., and Eelen, P. (2002). Affective priming for associatively unrelated primes and targets. *Psychol. Belg*. 42, 191–212. doi: 10.5334/pb.994

Hermans, D., Van den Broeck, A., and Eelen, P. (1998). Affective priming using a color-naming task: a test of an affective-motivational account of affective priming effects. *Z. Exp. Psychol*. 45:136–48.

Herring, D. R., White, K. R., Jabeen, L. N., Hinojos, M., Terrazas, G., Reyes, S. M., et al. (2013). On the automatic activation of attitudes: a quarter century of evaluative priming research. *Psychol. Bull*. 139, 1062–1089. doi: 10.1037/a0031309

Herring, D. R., White, K. R., Jabeen, L. N., Song, I., and Crites, S. L. (2015). Something (important) is out there! Effects of prime arousal and location on evaluative priming. *Motiv. Emot*. 39, 742–752. doi: 10.1007/s11031-015-9492-z

Hinojosa, J. A., Carretié, L., Méndez-Bértolo, C., Míguez, A., and Pozo, M. A. (2009). Arousal contributions to affective priming: electrophysiological correlates. *Emotion* 9, 164. doi: 10.1037/a0014680

Houghton, G., and Tipper, S. P. (1994). "A model of inhibitory mechanisms in selective attention," in *Inhibitory Processes in Attention, Memory, and Language*, eds D. Dagenbach and T. H. Carr (Academic Press), 53–112.

Hutchison, K. A. (2003). Is semantic priming due to association strength or feature overlap? A microanalytic review. *Psychon. Bull. Rev*. 10, 785–813. doi: 10.3758/BF03196544

Itkes, O., Kimchi, R., Haj-Ali, H., Shapiro, A., and Kron, A. (2017). Dissociating affective and semantic valence. *J. Exper. Psychol.: Gen.* 146, 924–942. doi: 10.1037/xge0000291

Itkes, O., and Kron, A. (2019). Affective and semantic representations of valence: a conceptual framework. *Emot. Rev*. 11, 283–293. doi: 10.1177/1754073919868759

Kawakami, N., and Miura, E. (2019). Tracking hand movements captures the response dynamics of the evaluative priming effect. *Cogn. Emot.* 33, 452–465. doi: 10.1080/02699931.2018.1483897

Kemp-Wheeler, S. M., and Hill, A. B. (1992). Semantic and emotional priming below objective detection threshold. *Cogn. Emot*. 6, 113–128.

Kiefer, M., and Martens, U. (2010). Attentional sensitization of unconscious cognition: Task sets modulate subsequent masked semantic priming. *J. Exp. Psychol. Gen*. 139, 464–489. doi: 10.1037/a0019561

Kiesel, A., Kunde, W., Pohl, C., and Hoffmann, J. (2006). Priming from novel masked stimuli depends on target set size. *Adv. Cogn. Psychol.* 2, 37–45. doi: 10.2478/v10053-008-0043-y

Klauer, K. C., Becker, M., and Spruyt, A. (2016). Evaluative priming in the pronunciation task. A preregistered replication and extension. *Exper. Psychol.* 63, 70–78. doi: 10.1027/1618-3169/a000286

Klauer, K. C., and Musch, J. (2001). Does sunshine prime loyal? Affective priming in the naming task. *Q. J. Exper. Psychol.* A54, 727–751.

Klauer, K. C., and Musch, J. (2002). Goal-dependent and goal-independent effects of irrelevant evaluations. *Pers. Social Psychol. Bull.* 28, 802–814. doi: 10.1177/0146167202289009

Klauer, K. C., and Musch, J. (2003). "Affective priming: Findings and theories," in *The Psychology of Evaluation: Affective Processes in Cognition and Emotion*, eds J. Musch and K. C. Klauer (Mahwah, NJ: Lawrence Erlbaum Associates Publishers), 7–49.

Klauer, K. C., Rossnagel, C., and Musch, J. (1997). List-context effects in evaluative priming. *J. Exp. Psychol. Learn. Mem. Cogn.* 23, 246. doi: 10.1037/0278-7393.23.1.246

Klauer, K. C., and Stern, E. (1992). How attitudes guide memory-based judgments: a two-process model. *J. Exp. Soc. Psychol.* 28, 186–206. doi: 10.1016/0022-1031(92)90038-L

Klauer, K. C., Teige-Mocigemba, S., and Spruyt, A. (2009). Contrast effects in spontaneous evaluations: a psychophysical account. *J. Pers. Soc. Psychol.* 96, 265–287. doi: 10.1037/a0013248

Klinger, M. R., Burton, P. C., and Pitts, G. S. (2000). Mechanisms of unconscious priming: I. response competition, not spreading activation. *J. Exp. Psychol. Learn. Mem. Cogn.* 26, 441. doi: 10.1037/0278-7393.26.2.441

Kumar, A. A., Steyvers, M., and Balota, D. A. (2021). A critical review of network-based and distributional approaches to semantic memory structure and processes. *Top. Cogn. Sci.* 14, 54–77. doi: 10.1111/tops.12548

Kunde, W., Kiesel, A., and Hoffmann, J. (2003). Conscious control over the content of unconscious cognition. *Cognition.* 88, 223–242. doi: 10.1016/S0010-0277(03)00023-4

Lai, V. T., Hagoort, P., and Casasanto, D. (2012). Affective primacy vs. cognitive primacy: Dissolving the debate. *Front. Psychol.* 3, 243. doi: 10.3389/fpsyg.2012.00243

Lemonnier, A., and Alexopoulos, T. (2019). The modulation of the evaluative priming effect by phasic affect. *Acta Psychologica* 192, 94–103. doi: 10.1016/j.actpsy.2018.11.001

Lucas, M. (2000). Semantic priming without association: a meta-analytic review. *Psychon. Bull. Rev.* 7, 618–630. doi: 10.3758/BF03212999

Maier, M. A., Berner, M. P., and Pekrun, R. (2003). Directonality of affective priming: effects of trait anxiety and activation level. *Exp. Psychol.* 50, 116. doi: 10.1026//1618-3169.50.2.116

Masson, M. E. J. (1995). A distributed memory model of semantic priming. *J. Exp. Psychol. Learn. Mem. Cogn.* 21, 3–23. doi: 10.1037/0278-7393.21.1.3

Mather, M., and Sutherland, M. R. (2011). Arousal-biased competition in perception and memory. *Perspect. Psychol. Sci.* 6, 114–133. doi: 10.1177/1745691611400234

McNamara, T. P. (2005). *Semantic Priming: Perspectives From Memory and Word Recognition*. New York, NY: Psychology Press.

McRae, K., De Sa, V. R., and Seidenberg, M. S. (1997). On the nature and scope of featural representations of word meaning. *J. Exp. Psychol. Gen.* 126, 99. doi: 10.1037/0096-3445.126.2.99

Moors, A. (2007). Can cognitive methods be used to study the unique aspect of emotion: an appraisal theorist's answer. *Cogn. Emot.* 21, 1238–1269. doi: 10.1080/02699930701438061

Moors, A., and De Houwer, J. (2006). Automaticity: a theoretical and conceptual analysis. *Psychol. Bull.* 132, 297. doi: 10.1037/0033-2909.132.2.297

Moors, A., and Scherer, K. R. (2013). "The role of appraisal in emotion," in *Handbook of Cognition and Emotion*, eds M. D. Robinson, E. Watkins, and E. Harmon-Jones (New York, NY: The Guilford Press), 135–155.

Murphy, S. T., and Zajonc, R. B. (1993). Affect, cognition, and awareness: affective priming with optimal and suboptimal stimulus exposures. *J. Pers. Soc. Psychol.* 64, 723. doi: 10.1037/0022-3514.64.5.723

Neely, J. H. (1991). Semantic priming effects in visual word recognition: a selective review of current findings and theories. *Basic Processes Reading Visual Word Recognition* 11, 264–336.

Neely, J. H., Keefe, D. E., and Ross, K. L. (1989). Semantic priming in the lexical decision task: roles of prospective prime-generated expectancies and retrospective semantic matching. *J. Exp. Psychol. Learn. Mem. Cogn.* 15, 1003. doi: 10.1037/0278-7393.15.6.1003

Neumann, R., and Lozo, L. (2012). Priming the activation of fear and disgust: evidence for semantic processing. *Emotion* 12, 223. doi: 10.1037/a0026500

Neumann, R., Schneider, L. J., and Kozlik, J. (2020). Goal-dependent automatic processing of potency: a potency-priming paradigm. *J. Exp. Soc. Psychol.* 89, 103975. doi: 10.1016/j.jesp.2020.103975

Nicolas, G., Fiske, S. T., Koch, A., Imhoff, R., Unkelbach, C., Terache, J., et al. (2022). Relational versus structural goals prioritize different social information. *J. Pers. Soc. Psychol.* 122, 659–682. doi: 10.1037/pspi0000366

Niedenthal, P. M., Rohmann, A., and Dalle, N. (2003). "What is primed by emotion concepts and emotion words?," in *The Psychology of Evaluation: Affective Processes in Cognition and Emotion*, eds J. Musch and K. C. Klauer (Mahwah, NJ: Lawrence Erlbaum Associates Publishers), 303–333.

Nummenmaa, L., Hyön,ä, J., and Calvo, M. G. (2010). Semantic categorization precedes affective evaluation of visual scenes. *J. Exp. Psychol. Gen.* 139, 222. doi: 10.1037/a0018858

Oosterhof, N. N., and Todorov, A. (2008). The functional basis of face evaluation. *Proc. Natl. Acad. Sci.* 105, 11087–11092. doi: 10.1073/pnas.0805664105

Osgood, C. E., Suci, G. J., and Tannenbaum, P. H. (1957). *The Measurement of Meaning (No. 47)*. Urbana, IL: University of Illinois Press.

Payne, B. K., Cheng, C. M., Govorun, O., and Stewart, B. D. (2005). An inkblot for attitudes: Affect misattribution as implicit measurement. *J. Pers. Soc. Psychol.* 89, 277–293. doi: 10.1037/0022-3514.89.3.277

Payne, K., and Lundberg, K. (2014). The affect misattribution procedure: Ten years of evidence on reliability, validity, and mechanisms. *Soc. Pers. Psychol. Compass.* 8, 672–686. doi: 10.1111/spc3.12148

Peeters, G. (1983). "Relational and informational patterns in social cognition." in Current Issues in European Social Psychology, Vol. 1, eds W. Doise and S. Moscovici (Cambridge, Cambridge University Press), 201–237.

Rohr, M., Degner, J., and Wentura, D. (2012). Masked emotional priming beyond global valence activations. *Cogn. Emot.* 26,224–244. doi: 10.1080/02699931.2011.576852

Rohr, M., Degner, J., and Wentura, D. (2015). The "emotion misattribution" procedure: Processing beyond good and bad under masked and unmasked presentation conditions. *Cogn. Emot.* 29, 196–219. doi: 10.1080/02699931.2014.898613

Rohr, M., Degner, J., and Wentura, D. (2020). Emotion misattribution from complex scene pictures: evidence for affective processing beyond valence. *Emotion*. doi: 10.1037/emo0000901 [Epub ahead of print].

Rohr, M., and Wagner, A. (2020). How monitor characteristics affect human perception in visual computer experiments: CRT vs. LCD monitors in millisecond precise timing research. *Sci. Rep.* 10, 1–10. doi: 10.1038/s41598-020-63853-4

Rohr, M., and Wentura, D. (2014). Spatial frequency filtered images reveal differences between masked and unmasked processing of emotional information. *Conscious. Cogn.* 29,141–158. doi: 10.1016/j.concog.2014.08.021

Rothermund, K., and Wentura, D. (1998). Ein fairer Test f?r die Aktivationsausbreitungshypothese: affektives Priming in der Stroop-Aufgabe. *Zeitschrift f?r Experimentelle Psychologie*. 45, 120–135.

Rothermund, K. K., and Werner, B. (2014). Affective priming in the valent/neutral categorisation task is due to affective matching, not encoding facilitation: reply to spruyt. *Cogn. Emotion* 28, 570–576. doi: 10.1080/02699931.2013.856778

Russell, J. A., and Barrett, L. F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: dissecting the elephant. *J. Pers. Soc. Psychol.* 76, 805–819. doi: 10.1037/0022-3514.76.5.805

Scherer, D., and Wentura, D. (2018). Combining the post-cue task and the perceptual identification task to assess parallel activation and mutual facilitation of related primes and targets. *Exp. Psychol.* 65, 84–97. doi: 10.1027/1618-3169/a000396

Scherer, D., and Wentura, D. (2022). Effects of evaluative homogeneity in working memory. [unpublished].
Scherer, K. R. (2005). What are emotions? And how can they be measured? *Soc. Sci. Inform.* 44, 695–729. doi: 10.1177/0539018405058216
Scherer, K. R., and Moors, A. (2019). The emotion process: event appraisal and component differentiation. *Annu. Rev. Psychol.* 70, 719–745. doi: 10.1146/annurev-psych-122216-011854
Schmitz, M., and Wentura, D. (2012). Evaluative priming of naming and semantic categorization responses revisited: a mutual facilitation explanation. *J. Exp. Psychol. Learn. Mem. Cogn.* 38, 984–1000. doi: 10.1037/a0026779
Schmitz, M., Wentura, D., and Brinkmann, T. (2014). Evaluative priming in a semantic flanker task: ERP evidence for a mutual facilitation explanation. *Cogn. Affect Behav. Neurosci.* 14, 426–442. doi: 10.3758/s13415-013-0206-2
Spruyt, A. (2014). Attention please: evaluative priming effects in a valent/non-valent categorisation task (reply to Werner and Rothermund, 2013). *Cogn. Emot.* 28, 560–569. doi: 10.1080/02699931.2013.832653
Spruyt, A., De Houwer, J., Everaert, T., and Hermans, D. (2012). Unconscious semantic activation depends on feature-specific attention allocation. *Cognition* 122, 91–95. doi: 10.1016/j.cognition.2011.08.017
Spruyt, A., De Houwer, J., and Hermans, D. (2009). Modulation of automatic semantic priming by feature-specific attention allocation. *J. Mem. Lang.* 61, 37–54. doi: 10.1016/j.jml.2009.03.004
Spruyt, A., De Houwer, J., Hermans, D., and Eelen, P. (2007). Affective priming of nonaffective semantic categorization responses. *Experim. Psychol.* 54, 44–53. doi: 10.1027/1618-3169.54.1.44
Spruyt, A., and Hermans, D. (2008). Affective priming of naming responses does not depend on stimulus repetition. *Can. J. Exp. Psychol.* 62, 237. doi: 10.1037/1196-1961.62.4.237
Spruyt, A., Hermans, D., Houwer, J. D., and Eelen, P. (2002). On the nature of the affective priming effect: affective priming of naming responses. *Soc. Cogn.* 20, 227–256. doi: 10.1521/soco.20.3.227.21106
Spruyt, A., Hermans, D., Pandelaere, M., De Houwer, J., and Eelen, P. (2004). On the replicability of the affective priming effect in the pronunciation task. *Experim. Psychol.* 51, 109–115. doi: 10.1027/1618-3169.51.2.109
Spruyt, A., and Tibboel, H. (2015). On the automaticity of the evaluative priming effect in the valent/non-valent categorization task. *PLoS ONE* 10, e0121564. doi: 10.1371/journal.pone.0121564
Spruyt, A., Tibboel, H., De Schryver, M., and De Houwer, J. (2018). Automatic stimulus evaluation depends on goal relevance. *Emotion* 18, 332. doi: 10.1037/emo0000361
Storbeck, J., and Clore, G. L. (2007). On the interdependence of cognition and emotion. *Cogn. Emot.* 21, 1212–1237. doi: 10.1080/02699930701 438020
Storbeck, J., and Clore, G. L. (2008). The affective regulation of cognitive priming. *Emotion* 8, 208. doi: 10.1037/1528-3542.8.2.208
Storbeck, J., and Robinson, M. D. (2004). Preferences and inferences in encoding visual objects: a systematic comparison of semantic and affective priming. *Pers. Soc. Psychol. Bull.* 30, 81–93. doi: 10.1177/0146167203258855
Topolinski, S., and Deutsch, R. (2013). Phasic affective modulation of semantic priming. *J. Exp. Psychol. Learn. Mem. Cogn.* 39, 414. doi: 10.1037/a002 8879
Usher, M., and Cohen, J. D. (1999). "Short term memory and selection processes in a frontal-lobe model," in *Connectionist Models in Cognitive Neuroscience. Perspectives in Neural Computing*, eds D. Heinke, G. W. Humphreys, and A. Olson (London: Springer). doi: 10.1007/978-1-4471-0813-9_7
VandenBos, G. R. (2007). *APA Dictionary of Psychology*. Washington, DC: American Psychological Association.
Warren, R. E. (1972). Stimulus encoding and memory. *J. Experim. Psychol.* 94, 90–100. doi: 10.1037/h0032786
Warren, R. E. (1974). Association, directionality, and stimulus encoding. *J. Experim. Psychol.* 102, 151–158. doi: 10.1037/h0035703
Wentura, D. (1998). *Affektives Priming in der Wortentscheidungsaufgabe: Evidenz f?r postlexikalische Urteilstendenzen [Affective priming in the lexical decision task: Evidence for post-lexical judgmental tendencies]. Sprache Kognition*. 17, 125–137.
Wentura, D. (1999). Activation and inhibition of affective information: Evidence for negative priming in the evaluation task. *Cogn. Emot.* 13, 65–91. doi: 10.1080/026999399379375
Wentura, D. (2000). Dissociative affective and associative priming effects in the lexical decision task: yes versus no responses to word targets reveal evaluative judgment tendencies. *J. Exp. Psychol. Learn. Mem. Cogn.* 26, 456. doi: 10.1037/0278-7393.26.2.456
Wentura, D., and Degner, J. (2010a). Automatic evaluation isn't that crude! Moderation of masked affective priming by type of valence. *Cognit. Emot.* 24, 609–628. doi: 10.1080/02699930902854587
Wentura, D., and Degner, J. (2010b). "A practical guide to sequential priming and related tasks," in *Handbook of Implicit Social Cognition: Measurement, Theory, and Applications*, eds B. Gawronski and B. K. Payne (Guilford), 95–116.
Wentura, D., and Frings, C. (2008). Response-bound primes diminish affective priming in the naming task. *Cogn. Emot.* 22, 374–384. doi: 10.1080/02699930701446064
Wentura, D., and Rohr, M. (2018). Emotion-specific priming effects with marginally perceptible facial expression primes: Evidence from the "leave-one-out" paradigm. *J. Experim. Psychol.* 44, 1946. doi: 10.1037/xhp0000581
Wentura, D., Rohr, M., and Degner, J. (2017). Masked emotional priming: a double dissociation between direct and indirect effects reveals non-conscious processing of emotional information beyond valence. *Conscious. Cogn.* 49, 203–214. doi: 10.1016/j.concog.2017.01.016
Wentura, D., and Rothermund, K. (2003). "The meddling-in of affective information: a general model ofautomatic evaluation effects," in *The Psychology of Evaluation: Affective Processes in Cognition and Emotion*, eds J. Musch and K. C. Klauer (Mahwah, NJ: Lawrence Erlbaum Associates Publishers), 51–86.
Werner, B., and Rothermund, K. (2013). Attention please: no affective priming effects in a valent/neutral-categorisation task. *Cogn. Emot.* 27, 119–132. doi: 10.1080/02699931.2012.711744
Werner, B., von Ramin, E., Spruyt, A., and Rothermund, K. (2018). Does sunshine prime loyal… or summer? Effects of associative relatedness on the evaluative priming effect in the valent/neutral categorisation task. *Cogn. Emot.* 32, 222–230. doi: 10.1080/02699931.2017.1282855
Zajonc, R. B. (1980). Feeling and thinking: preferences need no inferences. *Am. Psychol.* 35, 151. doi: 10.1037/0003-066X.35.2.151
Zhang, Q., Kong, L., and Jiang, Y. (2012). The interaction of arousal and valence in affective priming: behavioral and electrophysiological evidence. *Brain Res.* 1474, 60–72. doi: 10.1016/j.brainres.2012.07.023

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.

frontiers | Frontiers in Communication

ORIGINAL RESEARCH
published: 18 July 2022
doi: 10.3389/fcomm.2022.910023

Do People Get Used to Insulting Language?

Marijn E. Struiksma [1], Hannah N. M. De Mulder [2] and Jos J. A. Van Berkum [1]*

[1] Department of Languages, Literature and Communication, UiL OTS/Institute for Language Sciences, Utrecht University, Utrecht, Netherlands, [2] Leiden University Centre for Linguistics (LUCL), Leiden University, Leiden, Netherlands

Verbal insults go against a universal moral imperative not to inflict harm on others, and they also pose a threat to one's face or reputation. As such, these "verbal slaps in the face" provide a unique opportunity to explore the interface between language and emotion. We used electroencephalography (EEG) and skin conductance (SC) recordings to compare the short-term impact of verbal insults such as "Linda is an idiot" or "Paula is horrible" to that of more positive evaluations (e.g., "Linda is an angel", "Paula is impressive") and neutral factual descriptions (e.g., "Linda is a student"), examining how responses to these various speech acts adapt as a function of massive repetition. By using either the participant's or somebody else's name, we also explored how statement impact depended on who was being evaluated. Multilevel ERP analysis with three predesignated latency ranges revealed an early insult effect in P2 amplitude that was highly robust over repetition and also did not depend on who the insult was about. This P2 effect points to a very rapid and stable capture of emotional attention, plausibly triggered by the retrieval of evaluative word meaning from long-term memory. Insults also elicited a larger late positive potential (LPP), again regardless of who the insult was about, but this later effect did not withstand repetition. Skin conductance responses showed that insults did not lead to more arousal than compliments did. In all, our findings suggest that in a standard psycholinguistic comprehension experiment without real interaction between speakers, insults deliver lexical "mini-slaps in the face," such that the strongly negative evaluative words involved (e.g., "idiot") automatically grab attention during lexical retrieval, regardless of how often that retrieval occurs.

Keywords: psycholinguistics, communication, emotion, insults, morality, EEG, ERP, skin conductance

OPEN ACCESS

Edited by:
Francesca M. M. Citron, Lancaster University, United Kingdom

Reviewed by:
Luca Bischetti, University Institute of Higher Studies in Pavia, Italy
Cynthia Whissell, Laurentian University, Canada

****Correspondence:***
Marijn E. Struiksma
m.struiksma@uu.nl

Specialty section:
This article was submitted to Language Sciences, a section of the journal Frontiers in Communication

Received: *31 March 2022*
Accepted: *20 June 2022*
Published: *18 July 2022*

Citation:
Struiksma ME, De Mulder HNM and Van Berkum JJA (2022) Do People Get Used to Insulting Language? Front. Commun. 7:910023. doi: 10.3389/fcomm.2022.910023

INTRODUCTION

"Sticks and stones may break my bones, but words will never hurt me." Although consoling as an idea, we all know that this is not really true. Words can hurt, for at least one obvious reason: they are used to realize interpersonal behavior, and interpersonal behavior can sometimes really hurt. Think about receiving a break-up message, being criticized or insulted, or mocked. What is not so obvious, though, is the exact way in which words deliver their offensive, emotionally negative payload at the moment these words are being read or heard. Language comprehension is a highly complex multi-faceted process that rapidly activates or generates a variety of representations at various moments in time (e.g., Jackendoff, 2007; Tomasello, 2008; Van Berkum, 2018, 2019), such as the meaning of each of the words as retrieved from memory, the situation these words refer to, and the social intentions that can be attributed to the speaker.

This cascade of computed representations opens up the possibility that words can actually be offensive in multiple ways, related to different steps in the cascade, and at different moments in time.

Here, we explore the interface between words and their effects on people by examining the short-term impact of a particular type of offensive language, the verbal insult. One of the reasons why offensive expressions such as "you are ugly" or "you're a real asshole!" are frowned upon is because they violate a universal moral imperative to not inflict harm on others (Haidt, 2012; Greene, 2013; Gantman and Van Bavel, 2015). Despite this moral imperative, however, insults occur rather frequently in everyday life, in all corners of society, and with potentially serious consequences (Jay, 2009; Irvine, 2013). Understanding what an insulting expression does to people as it unfolds, and why, is therefore of considerable importance—to psycholinguists interested in how language moves people, but also to others who wish to understand the details of social behavior.

Everyday experience suggests that insults behave differently from more benign speech acts not only in their more negative "payload", but also in how easily they wear off. There are at least two sides to this. First, and sadly, many people seem to find it quite easy to brush off a compliment and forget about it as the day progresses, whereas a single insult may bother us for the rest of the day, and beyond. Second, whereas the impact of repeated compliments seems to adapt with repetition (just as many other good things do) or can in some cases even have detrimental effects (e.g., Brummelman et al., 2017), repeated insults do not seem to lose their sting. Of course, these are mostly informal observations. The main goal of our study is to examine whether evidence for the differential adaptation (or "saturation") to repeated verbal insults and repeated compliments can be observed in the lab, and, if so, which stage(s) of processing are implicated in the adaptation, and which are not. Such evidence can not only help us work out how insults and other moral transgressions are processed, initially and when repeated, but can also more generally advance our understanding of the complex interface between language and emotion. In passing, we may gain more insight into the everyday experience of why some things continue to move us, while others do not.

One relevant dimension is whether *you* or somebody else is insulted. Insults directed at you pose a severe threat to the self as well as to your reputation, and might for that reason not easily lose their sting—for members of an ultrasocial species that specializes in cooperation beyond the family (Leary, 2005; Tennie et al., 2010; Tomasello and Herrmann, 2010; Baumeister and Bushman, 2021), threats to one's reputation are not to be taken lightly. On the other hand, insults might also resist adaptation regardless of who the person evaluated is, because they inflict harm on others, because they are informative as to who is willing to do so, and because they signal a social conflict in your vicinity, possibly even in your group. Members of an ultrasocial species may well want to pay attention to such nearby verbal "slaps in the face" (Irvine, 2013).

On the reasonable assumption that verbal insults trigger a cascade of rapidly consecutive or overlapping processing effects, different parts of that cascade might be differently affected by repetition, with some of them rapidly wearing off, and others remaining strongly responsive for quite some time. In this study, we therefore use electroencephalography (EEG), a method that provides high temporal resolution as well as some qualitative indications of the processes involved.

Processing Nonverbal Emotional Stimuli

Emotional pictures and sounds attract attention and, as such, are given priority in subsequent processing (e.g., Vuilleumier, 2005; Schupp et al., 2006; Frijda, 2007; Kissler et al., 2009; Damasio, 2010; Panksepp and Biven, 2012; Pessoa, 2013; Carretié, 2014; Schupp and Kirmse, 2021). Many EEG studies have revealed such emotion-induced attention capture. In EEG research with emotional and neutral pictures, for example, emotional pictures can increase the P2, an anteriorly distributed short-lived ERP component around 200–250 ms after picture onset (Carretié et al., 2001; Delplanque et al., 2004; Huang and Luo, 2006). Emotional pictures have also been shown to modulate somewhat later EEG components, such as the early posterior negativity or EPN around 250–300 ms (Schupp et al., 2004, 2006; Herbert et al., 2008; Langeslag and Van Strien, 2018; Frank and Sabatinelli, 2019; Schupp and Kirmse, 2021), and the late positive potential or LPP, a deflection that typically emerges somewhere after 300 ms over large areas of the scalp, with a centro-parietal maximum, and a possibly lengthy duration (Schupp et al., 2004; Hajcak et al., 2013; Sabatinelli et al., 2013; Frank and Sabatinelli, 2019; Hajcak and Foti, 2020; Schupp and Kirmse, 2021).

As can be expected, negative and positive emotional pictures and sounds capture attention and are given priority in subsequent processing (Carretié, 2014). However, research on the negativity bias has revealed that people are on average particularly sensitive to negative events: such events not only recruit more attention and intensified processing than neutral events, but also often do so relative to positive events (Ito et al., 1998; Carretié et al., 2001; Rozin and Royzman, 2001; Huang and Luo, 2006; see Carretié, 2014, for a review). In EEG studies, for example, negative pictures have been reported to elicit a larger P2 (Carretié et al., 2001; Huang and Luo, 2006), a larger EPN (Schupp et al., 2004, 2006; Herbert et al., 2008), and a larger LPP (Ito et al., 1998; Huang and Luo, 2006; Foti et al., 2009) as compared to positive pictures. The exact source of the bias is currently under debate, with some arguing that it simply reflects statistical properties of the environment (Shin and Niv, 2021; Unkelbach et al., 2021), and others proposing an evolutionary analysis involving the degree to which negative vs. positive stimuli affect fitness (Lazarus, 2021). Also, the negativity bias does not guarantee that *every* negative stimulus or stimulus set captures more attention than every positive stimulus or stimulus set (Carretié, 2014). After all, a snapping shoelace is a lot less evocative than the birth of one's child. The negativity bias is real, but it exists as an average phenomenon, emerging for reasons that remain to be fully explained.

Processing Emotional Language

As might be expected, similar mechanisms of attention capture and subsequent intensified processing are at work when people read or listen to emotional language (e.g., Bernat et al., 2001;

Kissler et al., 2006, 2009; Kanske and Kotz, 2007; Herbert et al., 2008; Scott et al., 2009; Lei et al., 2017; see Citron, 2012 for a review). For example, relative to neutral words, negative (and positive) words have been reported to elicit an enhanced P2 (Herbert et al., 2006; Trauer et al., 2012; Wang and Bastiaansen, 2014), an enhanced EPN (e.g., Herbert et al., 2008; Recio et al., 2014; Rohr and Abdel Rahman, 2015), an enhanced N400 (e.g., Kanske and Kotz, 2007; Rohr and Abdel Rahman, 2015, for negative words only), as well as a larger LPP (Carretié et al., 2008; Herbert et al., 2008; Kissler et al., 2009; Wang and Bastiaansen, 2014).[1] Comparable emotion-induced EEG effects can be observed for words that are part of a larger discourse (e.g., Holt et al., 2009; Fields and Kuperberg, 2012, 2015, 2016). Although comparisons of emotional to neutral words often involve both valence and the degree of arousal (sympathetic system activation) that the word brings about, there is evidence that either one can play a role in the effects obtained, and that the two factors can also interact (see, e.g., Bayer et al., 2010; Citron et al., 2013; Recio et al., 2014; Espuny et al., 2018; Vieitez et al., 2021).

In addition to being a negative event, a verbal insult is also a morally objectionable statement, which makes EEG studies on the processing of morally objectionable language relevant to current concerns. In an early study exploring the impact of morally objectionable language with ERPs (Van Berkum et al., 2009), participants read statements that did or did not clash with their moral value system. Relative to acceptable words, words that rendered the unfolding statement morally objectionable for the participant at hand elicited an early P2-like positivity around 200-250 ms, and an increased LPP between 500 and 650 ms. These words also elicited a small centro-parietally distributed N400 effect, a well-known ERP effect that within the language domain is assumed to index some context-induced difficulty in retrieving and/or otherwise processing the meaning of a word (e.g., Kutas and Federmeier, 2000, 2011; Van Berkum, 2009; Brouwer et al., 2017). These various ERP effects have also been observed in other moral transgression studies (Foucart et al., 2015; Leuthold et al., 2015; Peng et al., 2017; Kunkel et al., 2018; Weimer et al., 2019), albeit not always all together. To the extent that they reappear in the current study, they can be useful in tracking the impact of a verbal transgression as a function of repeating it.

A few EEG studies have directly examined the processing of insulting language. In studies using isolated words, insulting words have been reported to elicit LPP effects (Carretié et al., 2008, relative to neutral but not complimenting words), occipital P1 effects (Wabnitz et al., 2012, relative to both), and increased mid-latency frontal negativities (Wabnitz et al., 2012, between 450 and 580 ms relative to neutral but not complimenting words; Wabnitz et al., 2016, between 395-480 ms, relative to both). EEG research on richer and more realistic insults is actually very rare, but in two recent studies with somewhat more contextualized insults, insulting words elicited a centro-parietally focused LPP effect between 500 and 800 ms in one study (Rohr and Abdel Rahman, 2018, relative to neutral but not complimenting words), and an increased anterior negativity between 300 and 400 ms followed by a small LPP effect between 600 and 900 ms in the other study (Otten et al., 2017, relative to compliments).

Although relative to compliments, insults do not always increase the amplitude of potentially relevant components, such as the P1, P2, EPN, N400 and LPP (e.g., Benau and Atchley, 2020) and certainly do not boost these components simultaneously in every study, the relevant EEG studies conducted so far do suggest that insulting language does typically capture attention, and elicits intensified processing very rapidly, in a way that is comparable to the impact of other verbal moral transgressions (e.g., Van Berkum et al., 2009). Behavioral research on verbal insults confirms this. In a lexical decision task, for example, word-pseudoword decisions are slower and more error-prone with insult words than with compliment words (Carretié et al., 2008). Furthermore, in emotional Stroop tasks, people are slower to name the color of insult words than of compliment words (Siakaluk et al., 2011). In all, and as might be expected from the negativity bias literature, verbal insults command attention to a larger extent than verbal compliments do, presumably due to their offensive nature.

What Exactly Is Offensive About an Insult?

What aspects of a verbal insult can give rise to such effects? Imagine hearing "you are an idiot!", uttered toward you in a way that is genuinely contemptuous. A recent model of affective language comprehension, the ALC model (Van Berkum, 2018, 2019; see **Figure 1**), predicts that multiple aspects of this verbal event can have an emotional impact. Most obviously, the speaker realizes a social intention (cf. Tomasello, 2008; Enfield, 2013), which in this case presumably is to overtly express his or her contempt for you, or at least to strike out at you, and possibly cause some hurt. That social act or "move" will usually elicit strong emotion—as part of our evolved cooperation-oriented way of living, human beings are all striving for respect (MacDonald and Leary, 2005; Greene, 2013), and an insult is a serious threat to one's "face" (Goffman, 1967; Brown and Levinson, 1987). Related to this, people want to "belong" and be close to at least some others, and a genuine insult signals strong interpersonal distancing—depending on who is speaking, that might hurt a bit, or a lot. At the level of social moves, genuine insults are very unpleasant things to receive.

The ALC model, however, predicts that an unfolding utterance such as "you are an idiot!" can generate at least three other emotionally evocative representations, that, although related to the social move, are not identical to it. One is the inferred referential intention of the speaker (cf. Tomasello, 2008): while the situation described by the speaker (you being an idiot) need not be a correct characterization of the state of affairs in the world, imagining this state as an unavoidable part of processing the utterance may nevertheless still elicit some emotion. A second possible trigger is the speaker's stance: a contemptuously uttered "you are an idiot!" will inevitably lead the listener to register contempt, regardless of what the exact target of that contempt

[1] The relationship of the LPP to various other late and roughly similarly distributed positive ERP deflections, such as the P3, P600, and Late Positive Complex (LPC) is currently under debate (e.g., Hajcak and Foti, 2020). Starting from emotion-oriented ERP research, we focus on LPP-relevant studies here, but we'll briefly return to the relationship with the P3 in the discussion section.

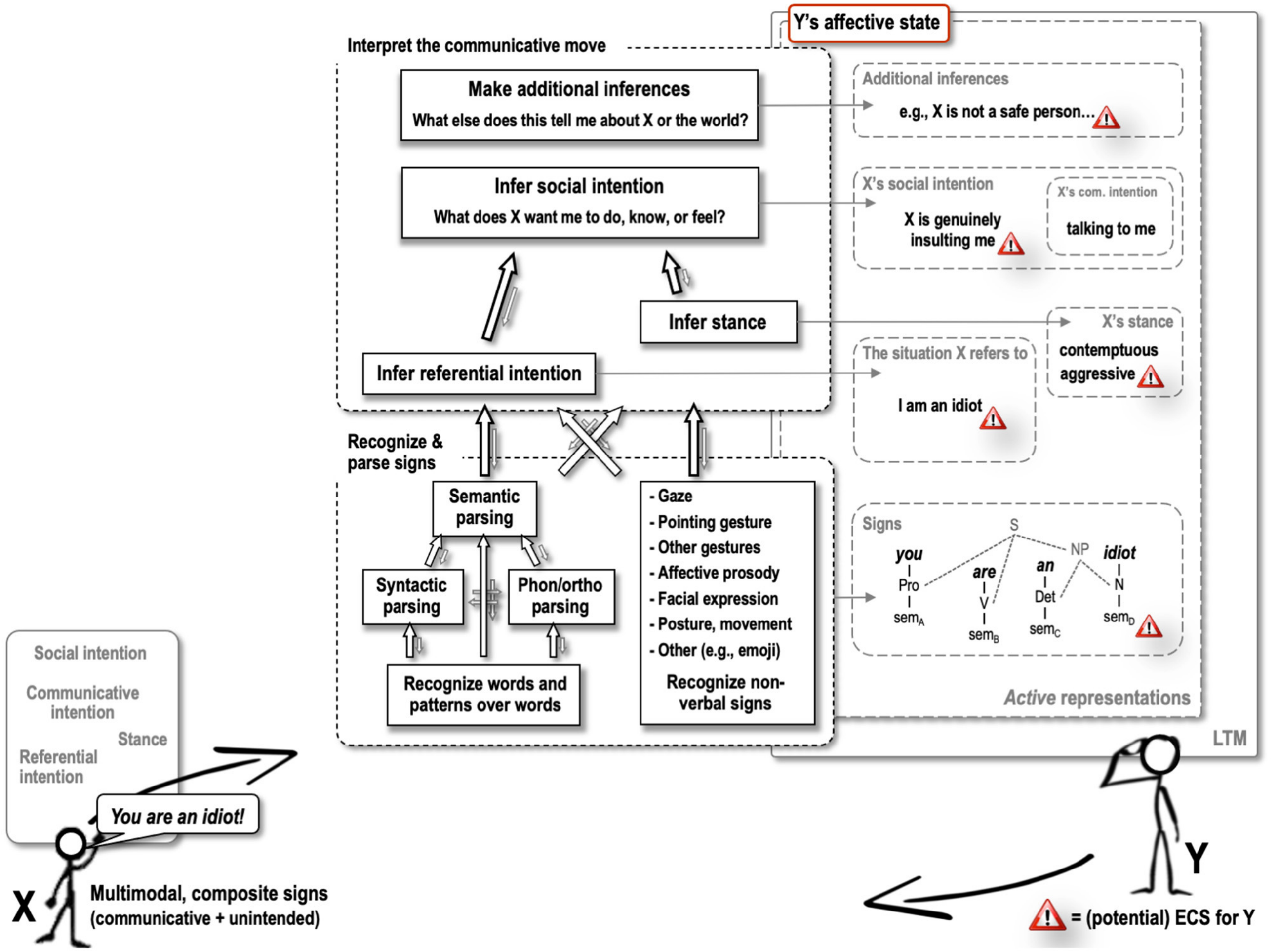

FIGURE 1 | The Affective Language Comprehension (ALC) model illustrated with an insult example. ECS, emotionally competent stimulus; LTM, long-term memory; Phon/ortho parsing, phonological or orthographic parsing; X's com. intention, X's communicative intention. See text and Van Berkum (2018, 2019) for details.

is, and this can be evocative in a relatively unspecific way. Finally, the basic signs that make up a more complex utterance can carry their own emotional "payload" (Van Berkum, 2018). For example, if particular words such as "idiot," "bitch" or "disappointment" have often been used to implement negative social moves, to express negative stances, or to describe very negative situations, or if those words have often been witnessed to elicit strong negative responses in others, simple emotional conditioning will lay down negative affective connotations as part of the memory trace for that particular word, sometimes even to the extent that the word becomes taboo (Jay, 2009). As with other emotionally conditioned stimuli (e.g., a light that sufficiently reliably predicts a shock), the mere use of the word can then evoke a rapid automatic affective response, independent of the precise sentence-level message and the social intention of the speaker on this particular occasion.

If you hear an insult directed at somebody else, the ALC model leads us to expect that things will be partly different, but also partly comparable. If the person evaluated is somebody you care for, the referential intention and associated social move may cause you to feel hurt or angry too, but if he or she is somebody you dislike, it might actually in part evoke a very different emotion (Singer et al., 2006). Other things might be less dependent on who the person evaluated is. For a species heavily invested in cooperation, for example, displays of an aggressive stance (such as a verbal or physical slap in the face) may automatically trigger negative emotion in the target of that aggression, as well as in those who witness somebody else being a target of aggression. Furthermore, and related, due to emotional conditioning, offensive words like "idiot," or "bitch" may elicit the same automatic affective response independent of who is being targeted, and what the word is used for in this specific communicative exchange.

The Experiment

The goal of our study is to use EEG to examine whether verbal insults are less sensitive to repetition than compliments are, and, if so, which stage(s) of processing are implicated in the adaptation, and which are not. In line with longstanding practices in EEG research, different EEG studies have used different latency ranges (and scalp regions) to quantify specific ERP components, possibly in part driven by the data. Recent work (Luck and Gaspelin, 2017) has drawn attention to the risks that are inherent to this flexibility in EEG analysis procedures. We therefore committed ourselves to only testing hypotheses in three predesignated critical latency ranges (and scalp regions) where, in our earlier EEG work (Van Berkum et al., 2009), we observed ERP effects of morally objectionable language: an "early" 200–250 ms latency range aimed at detecting P2 effects over the anterior part of the scalp, a "middle" 350–450 ms latency range aimed at detecting N400 effects over the central-posterior part of the scalp, and a "late" 500–650 ms latency range aimed at detecting LPP effects over the same central-posterior scalp region.

In line with the extant ERP literature, we made the following assumptions. First, insult-induced modulations of the word-elicited N400 component are assumed to index the impact of emotion on context-dependent lexical retrieval (or processes closely associated with that retrieval, see, e.g., Kutas and Federmeier, 2000, 2011; Lau et al., 2008; Van Berkum, 2009, 2012; Brouwer et al., 2017), involving usually left-dominant temporal as well as inferior frontal cortical generators (Halgren et al., 2002; Lau et al., 2008; Swaab et al., 2012). Second, an insult-induced increase of the preceding P2 will be taken as suggesting that emotional attention is up-regulating some aspect of visual stimulus processing (Delplanque et al., 2004; Huang and Luo, 2006; Hajcak et al., 2012; Carretié, 2014). Finally, an insult-induced increase of the LPP will be assumed to reflect additional—and possibly lengthy—elaborative processing of motivationally important, "significant" stimuli (e.g., Cacioppo et al., 1993, 1994, 2004; Schupp et al., 2000, 2004; Smith et al., 2003; Kisley et al., 2007; Sabatinelli et al., 2007, 2013; Holt et al., 2009; Frank and Sabatinelli, 2019; Hajcak and Foti, 2020), involving the extrastriate visual cortex as well as emotion-dedicated cortical areas (Frank and Sabatinelli, 2019), and possibly involving the activity of approach and avoidance systems (Bradley, 2009; Hajcak and Foti, 2020). Although it is tempting to assume that a "later" ERP effect like the LPP indexes later psycholinguistic computations as delineated in the ALC model (Van Berkum, 2019), early processing, such as the lexical retrieval of a word with strong emotional valence, may itself trigger additional processing in response to a motivationally relevant stimulus (i.e., a later LPP effect), regardless of whatever else happens in the psycholinguistic processing cascade laid out in the model.

In the experiment, participants read a series of repeated statements that realized three different speech acts: (1) insults, interpersonally evaluative statements that combined a person's name with a negative evaluative predicate (e.g., "Linda is horrible," "Paula is a liar"), (2) compliments, interpersonally evaluative statements that combined a person's name with a positive evaluative predicate (e.g., "Linda is impressive," or "Paula is an angel"), and (3) non-evaluative, factually correct descriptive statements (e.g., "Linda is Dutch," or "Paula is a student"). To examine whether the impact of the speech act depended on who the statement was about, half of these three sets of statements used the participant's own name, and the other half used the name of somebody else. Although the study did not involve real interaction, we presented the statements as "being uttered by three specific men, about the participant herself, or some other participant in the study."

Our main goal was to look for signs of differential adaptation to repeated insults and repeated compliments, and to explore which stages of language-induced processing are implicated in any differential adaptation, and which are not. To examine this, we exposed our participants to 90 insults featuring their name as well as 90 insults featuring another person's name, and compared the ERPs elicited by those two types of insults (in the abovementioned P2, N400, and LPP latency ranges) to the ERPs elicited by 90 compliments featuring their name and 90 compliments featuring the other person's name, all as a function of repetition. Apart from the fact that the act of insulting was repeated many times, each particular statement item, e.g., "Linda is horrible," or "Paula is an angel," was presented many times throughout the session as well. Statements were distributed over several blocks, and we used the development of ERP amplitudes across these blocks as a measure of how participants adapted to repetition. Furthermore, we included 90 factually correct descriptive statements with the participant's name, and 90 more with the other person's name, to diversify the materials, and to also have an indication of how the response to non-evaluative and hence presumably "neutral" statements developed over blocks.

Apart from capturing attention, negative as well as positive emotional stimuli can also lead to arousal, that is, increased sympathetic activity within the autonomic nervous system that is intended to prepare the body for action as well as possible damage (Lang et al., 1993; Codispoti et al., 2001; Dawson et al., 2007). Arousal can be measured via the phasic skin conductance (SC) response, a measure that picks up on short-lived increases in sweat production (Lang et al., 1993; Codispoti et al., 2001; Dawson et al., 2007; Boucsein, 2012), and that has been shown to detect increases in arousal to reprimands and taboo words (Harris et al., 2003; Harris, 2004; Eilola and Havelka, 2011). In our study, skin conductance recording allowed us to assess the arousing potential of insults and compliments, again as a function of who was involved, and as a function of massive repetition.

Our predictions were as follows.

1. Based on our earlier EEG work with morally offensive language (Van Berkum et al., 2009), we expected verbal insults to generate a larger P2, a larger N400, and a larger LPP, all relative to compliments, but presumably also relative to neutral statements.
2. The massive repetition of evaluative statements should influence at least some of these three ERP responses to insults less than the corresponding ERP responses to compliments. So, for example, even though the amplitude of the insult-induced P2 might decrease over repetitions, the amplitude of the compliment-induced P2 should decrease over repetitions more, due to a stronger degree of adaptation to compliments. We had no strong predictions as to whether this attenuated adaptation for insults (relative to compliments) would involve the P2, the N400, or the LPP.

3. The differential impact of insults (relative to compliments, as well as neutral statements) should be larger for speech acts that described the participant than for speech acts that described somebody else, again in the P2, the N400, and/or the LPP.
4. We considered it likely that the differential adaptation to repetition for insults and compliments would depend on who the statement was about. In particular, we expected participants to adapt least to insults about themselves, and made no predictions about the other three cases.
5. With respect to skin conductance, the threat to the self that is posed by insults featuring the participant's own name led us to expect the strongest arousal (i.e., the highest phasic SC response) here, and we also expected these responses to be the most resistant to repetition. We had no strong predictions for the SC response to the other speech act types. Strongly positive stimuli are known to generate arousal too (e.g., Lang et al., 1993), so compliments to participants might lead to more arousal than compliments to somebody else, and neutral factual statements. However, if participants feel threatened by compliments to others, for example, a different arousal profile might arise.

MATERIALS AND METHODS

Participants

We tested 79 right-handed Dutch participants, all students at Utrecht University who identified themselves as female[2] (M_{age} = 20.90, $Range_{age}$ = 18–30). They were non-dyslexic and did not report any neurological disorders. Participants who had a partner named "Bram," "Daan" or "Paul" could not sign up for the experiment, since those were the names of the men that supposedly uttered the statements. All research was conducted in line with the WMA Declaration of Helsinki—Ethical Principles for Medical Research Involving Human Subjects, as well as The Netherlands Code of Conduct for Scientific Practice issued in 2004 (revised in 2012) by the Association of Universities in the Netherlands (VSNU). The study was approved by the local Ethics Assessment Committee Linguistics (case number and relevant documents available upon request). Before the experiment, participants read a detailed information letter and accompanying informed consent form. These carefully explained the procedure and the offensive nature of some of the stimuli. We also made explicit, both in the consent form and face-to-face at the start of each session, that participants could decide to terminate the experiment at any time without having to provide a reason, and still receive the full financial compensation (10,- euro/hour). Before the experiment started the participant signed the informed consent form. None of the participants resigned.

Materials and Design

To realize our Speech Act (insult, compliment, neutral description) by Person Described (participant, other) design, we created 10 unique insult frames (e.g., "<X> is horrible," "<X> is a liar"), 10 unique compliment frames (e.g., "<X> is impressive," or "<X> is an angel"), and 10 unique non-evaluative, factually correct descriptive statements (e.g., "<X> is Dutch," or "<X> is a student"), and combined each with the participant's name or somebody else's name just before the experiment. Critical words, always in sentence-final position, were selected on the basis of a written pretest with 42 female Dutch students from Utrecht University (M_{age} = 21.07, $Range_{age}$ = 18–28) who did not participate in the main experiment. Pretest participants read a large set of insulting, complimenting and non-evaluative neutral words embedded in a larger "<female-name> is <evaluative-predicate>" phrase, and rated the phrases on a seven-point scale ranging from "very insulting" (−3) to "neutral" (0) to "very complimenting" (3). The ratings were used to choose 10 strongly insulting words [e.g., "horrible," "liar," $M_{insults}$ = −2.39 (± 0.23)] as well as 10 equally strongly complimenting words [e.g., "impressive," "angel," $M_{compliments}$ = 2.45 (±0.18)], and 10 neutral words [e.g., "Dutch," "student," $M_{neutrals}$ = 0.05 (±0.10)] with strength expressed as deviation from the scale's neutral midpoint. The absolute deviation from the scale's neutral midpoint did not differ for insulting and complimenting word sets [$t_{(18)} = 0.64$, $p = 0.53$]. The word sets were also matched on length [$M_{insults}$ = 9.00 (±4.35), $M_{compliments}$ = 9.00 (±2.16), $M_{neutrals}$ = 9.00 (±3.16); $F_{(2,27)} = 0.00$, $p = 1.00$] and frequency [$M_{insults}$ = 49.45 (±78.91), $M_{compliments}$ = 55.28 (±56.08), $M_{neutrals}$[3] = 125.95 (±255.30); $F_{(2,27)} = 0.73$, $p = 0.49$] according to the SUBTLEX-frequency database for Dutch (Keuleers et al., 2010). In each set, half of the trials contained an indefinite article ("een" in Dutch), the other half did not. The full set of stimuli can be found in **Supplementary Table S1**.

Every trial started with a centered fixation-cross and a black screen paired with a beep, followed by a word-by-word presentation of the sentence. Each word was presented in the center of the screen in Calibri font size 35, with a duration that varied with word length (imitating natural reading parameters, see Otten and Van Berkum, 2007), and with the critical sentence-final (insulting, complimenting, or neutral) word shown for 1,000 ms in Calibri font size 50. After the critical word, another black screen was presented.

Statements were presented in blocks, with separate blocks for each of the Speech Act by Person Described combinations: insults involving the participant, compliments involving the participant, neutral observations involving the participant, insults involving somebody else, compliments involving somebody else, and neutral observations involving somebody else. In each block, each of the 10 relevant statements (e.g., 10 specific insults about the participant, such as "Linda is horrible") was presented three times, resulting in 30 stimuli per block. Within each block the 30 stimuli were arranged in pseudorandom order such that a specific insult was never presented twice in a row. Each of these six 30-item blocks was presented three times in a different pseudorandomized order, resulting in 18 blocks per session, and nine occurrences of any critical statement (e.g., "Linda is horrible," or "Paula is impressive") across the entire session. The order of the 18 blocks was counterbalanced according to a Latin

[2]We limited ourselves to a single gender to facilitate the construction of materials.

[3]Due to experimenter error, the frequency of one of the neutral words ("woman") was incorrectly coded, leading to a higher mean frequency of the neutral words set. Relative to the standard deviation across items, however, the mean difference was fairly small (as indicated by a nonsignificant test).

square procedure, resulting in six comparable stimulus lists. Half of these started with a statement block involving the participant's name, the other half with a statement block involving the length-matched name of somebody else. Orthogonal to this, one-third started with insults, one-third with compliments, and one-third with a neutral block. In all, each session contained 18 blocks of 30 statements each, which together lasted ∼70 min. The effect of statement repetition was assessed *via* this Block (1–18) factor.

As they worked through each statement block, participants had no other task than to read each statement attentively. To keep participants sufficiently engaged, we also presented a very short auditory oddball task after each statement block, in which they had to detect 20 high beep targets (2,000 Hz) among 80 low beep distractors (1,000 Hz) and press a button when a target was detected as fast as possible. Each oddball minitask took 50 s.

Procedure

After asking the participant for informed consent and explaining face-to-face that they could resign at any time without losing their compensation, the EEG and skin conductance (SC) electrodes were applied. Statement stimuli were presented using Neurobehavioral Systems' "Presentation" software, on a 22″ Iiyaman Prolite screen with a display resolution of 1,920 × 1,080 pixels and a 60 Hz refresh rate. Text was presented in Calibri font size 35, except for the critical word which was presented in font size 50, color 180, 180, 180 RGB. The distance between the screen and the participant was ∼70 cm. At the start of the experiment the participant was introduced to four fictional characters: "In this experiment there are three men who will give some remarks, their names are Bram, Daan and Paul. Read their remarks carefully. The remarks will be about you or another woman. The other woman is a student in Utrecht and a participant in this experiment as well." This explanation was followed by six practice trials, which used evaluative words that were not in the experimental stimulus set. After the statement practice trials, the oddball task was explained and briefly practiced as well.

In the main session, evaluative or neutral factual statement blocks alternated with auditory oddball blocks. Prior to each statement block we mentioned which man the statements would come from in this block: always Bram for the neutral statements, and always Daan and Paul for the insults and compliments. We counterbalanced the latter across sessions such that a single man was presented as the source of insults to the participant and compliments to the other woman, and the other as the source of compliments to the participant and insults to the other woman (so that these men could be experienced as discriminative, rather than as, e.g., "always insulting everybody"). Blocks of statements and oddball were separated by a short break, and after six blocks of each task there was an obligatory longer break during which the experimenter checked how the participant was doing. Finally, the participant was asked to fill out an exit questionnaire and several personality questionnaires [the State-Trait Anxiety Inventory (STAI, Spielberger et al., 1983), the Rosenberg Self-Esteem Scale (RSES, Rosenberg, 1965), the Interpersonal Reactivity Index (IRI, Davis, 1983), the 10-item Connor-Davidson Resilience Scale (CD-RISC 10, Campbell-Sills and Stein, 2007), and the dimensions Outlook and Resilience (Davidson and Begley, 2013)]. The questionnaires were added to investigate potential individual differences in an exploratory fashion. The entire experiment lasted ∼2 h.

EEG and Skin Conductance Measurements

EEG was recorded using 64 Ag/AgCl electrode BioSemi caps with a 10–20 configuration. We used two additional electrodes at the left and right mastoid for re-referencing during the analysis, two facial electrodes at the left and right outer canthi to measure the horizontal electrooculogram (hEOG), and another two electrodes above and below the left eye to measure vertical EOG (vEOG), in order to detect blinks and eye movements. Skin conductance was measured with two stainless-steel Nihon Kohden electrodes at the distal phalanx of the index finger and middle finger of the left hand. A BioSemi ActiveTwo system was used to sample and record all EEG and SC data at 2,048 Hz.

EEG Analysis

The EEG data was preprocessed and analyzed offline using BrainVision Analyzer 2.1 software (Brain Products), Matlab (R2019a) and Rstudio (1.2.5042, R version 3.6.3). First, EEG signals were re-referenced to the average of the left and right mastoid, bandpass-filtered at 0.1–35 Hz (24 dB slope), and downsampled to 500 Hz. Next, all markers were checked using a customized macro, and segment-relevant condition markers were inserted at the onset of the critical word (CW). Next, the data was segmented into epochs from 200 ms before to 1,000 ms after CW onset, and segments were baseline-corrected by subtracting the mean signal amplitude in the CW-preceding 200 ms from all individual amplitude values in the remaining 1,000 ms. To optimize the signal-to-noise ratio, subsequent artifact rejection was done in two steps that focused on EOG and data channels, respectively. First, to remove eye movements and blinks, all epochs in which the bipolar hEOG and/or vEOG signal exceeded ±75 μV, or which displayed a voltage step of 50 μV or more between two neighboring sampling points, or in which the difference in signal activity was lower than 0.5 μV in an interval of 100 ms, were marked as bad trials. Second, the 64 regular EEG-channels were tested with the same criteria, but now only the specific channels failing the test were marked to be excluded from further analysis. As a final step of the preprocessing of the data the segments were exported to MAT-files for further processing in Matlab.

In Matlab all segments and channels within segments marked for exclusion were removed. Based on prior EEG work with verbal moral transgressions in our lab (Van Berkum et al., 2009), we computed mean amplitudes per segment and electrode for three predesignated latency-of-interest ranges that had also been used to assess P2, N400 and LPP effects in that earlier study: an "early" 200–250 ms latency range (optimized for detecting P2 effects), a "middle" 350–450 ms latency range (optimized for detecting N400 effects), and a "late" 500–650 ms latency range (optimized for detecting LPP effects)[4]. Furthermore, and

[4]As a compromise between strict replication and compatibility with the wider literature, we widened the rather narrow N400 latency range used in Van Berkum et al. (2009) with an extra 50 ms, while keeping it centered on the 400 ms mark (so from 375–425 to 350–450 ms).

again based on our earlier verbal moral transgressions study, we optimized our sensitivity for detecting any P2, N400 or LPP effects by computing, per segment, an average EEG amplitude over the anterior part of the brain (involving electrodes Fp1, AF3, AF7, F1, F3, F5, F7, FC1, FC3, FC5, FT7, Fp2, AF4, AF8, F2, F4, F6, F8, FC2, FC4, FC6, and FT8) for the 200–250 ms latency range, and an average EEG amplitude over the central and posterior part of the brain (involving electrodes C1, C3, C5, T7, CP1, CP3, CP5, TP7, P1, P3, P5, P7, PO3, PO7, O1, C2, C4, C6, T8, CP2, CP4, CP6, TP8, P2, P4, P6, P8, PO4, PO8, O2) for the 350–450 ms latency range as well as for the 500-650 ms latency range. These predesignated latency- and region-of-interest data operations condensed each of the participant's EEG signal in a critical epoch to just three hypothesis-relevant amplitude data points (anterior 200–250 ms, posterior 350–450 ms, and posterior 500–650 ms), all exported into a text file for further analysis using Rstudio.

The three latency ranges were analyzed separately in Rstudio using linear mixed models. To further reduce noise, we first averaged the three repetitions of each specific critical item (e.g., "Linda is horrible") within each block, such that three repetition averages remained (one for each relevant block). The participant and statement variances were included as random factors and were estimated simultaneously, resulting in a cross-classified model (Quené and van den Bergh, 2004, 2008). For each latency range, we constructed models for the mean amplitude by iteratively adding potentially relevant components and testing for significant model improvement at each addition (using the likelihood ratio (−2LL difference chi-square) test, $p < 0.05$). We started by comparing an empty model, containing only the random factors for participant and statement, to the base model containing the random factors and the three main effects of interest: Speech Act (insult, neutral, compliment), Person Described (you, other) and Block (1–18). The factor Block was centered to avoid correlation between the linear and quadratic effect that was added in the next step. We then added another factor in each next step, starting with the 2-way interactions and working all the way up to the 4-way interaction. The final model was in fact the full factorial model. We then evaluated which model fit the data best and only kept those factors that explained a significant amount of variance once they were added to the model or which were necessary to test hypothesized interactions (Winter, 2019). Factors that did not significantly improve the model were dropped (H. van den Bergh, personal communication; Seasholtz and Kowalski, 1993). The summary of arriving at the best model is given in **Table 1**. The best and most sparse model was used to estimate the model parameters correctly. For transparency, the full factorial models are also reported in **Supplementary Tables S2–S4**. For the *post-hoc* analyses we used Tukey HSD corrected pairwise comparisons.

Figures with results from the linear mixed model analysis, reporting the mean amplitudes for each latency range, were based on parameter estimates from the best models, generated using Rstudio. The ERP and topography results were generated in BrainVision Analyser. For visualization the data were processed from the segmentation step as follows. The data were segmented into the 18 blocks of the experiment. For each block the data was then segmented into epochs from 200 ms before to 1,000 ms after the onset of every critical word, and segments were baseline-corrected by subtracting the mean signal amplitude in the first 200 ms of each epoch from all individual amplitude values in that epoch. Next, the artifact rejection procedure based on the EOG-channels was performed with similar settings as mentioned above. However, instead of only marking the bad trials, they were removed from the analysis right away. This was also done for the artifact rejection based on the remaining 64 electrodes. The channels that were marked as bad were immediately excluded from further analysis. As a final step, the segments were averaged per block. For visualization purposes new aggregate channels representing four quadrants were calculated by averaging the following channels. A Left Anterior (LA) channel was calculated from electrodes Fp1, AF3, AF7, F1, F3, F5, F7, FC1, FC3, FC5, and FT7; a Left central/Posterior (LP) channel was calculated from electrodes C1, C3, C5, T7, CP1, CP3, CP5, TP7, P1, P3, P5, P7, PO3, PO7, and O1; a Right Anterior (RA) channel was calculated from electrodes Fp2, AF4, AF8, F2, F4, F6, F8, FC2, FC4, FC6, and FT8; and a Right Central/Posterior (RP) channel was calculated from electrodes C2, C4, C6, T8, CP2, CP4, CP6, TP8, P2, P4, P6, P8, PO4, PO8, and O_2. The average response per Speech Act by Person Described and by number of repetition (1–3) was computed for these four channels.

SC Analysis

The SC data was preprocessed offline using BrainVision Analyzer 2.1 software (Brain Products). The data was downsampled to 500 Hz and segmented into epochs from the onset until 165 s after each Speech Act by Person Described block, this epoch included the entire block of 30 trials. These epochs were exported into mat-files and further analyzed using Matlab R2019b and the Ledalab toolbox (Benedek and Kaernbach, 2010). We used the Continuous Decomposition Analysis to decompose the phasic and tonic activity. The data was down sampled to 10 Hz and adaptive smoothing using a Gaussian window was applied. Six sets of initial values were used to optimize the solution. The event-related phasic activity was further analyzed. Since the final 30th trial in each block did not always have 4 s of clean signal after critical word onset, which was necessary for further analysis, we only analyzed the first 29 trials. We used a 10 μS threshold for significant phasic activity and computed the mean activity in the time window of 1–4 s after the critical event as the area under the curve (squared integrated skin conductance response: $ISCR^2$). The results were log transformed to normalize the data, and exported to a text-file for further analysis using Rstudio. The linear mixed model analysis in Rstudio followed the same procedure as reported above for the EEG analysis: the models were built iteratively until the full-factorial model was reached and subsequently the most sparse model was achieved by removing non-significant factors that did not improve the model fit. This procedure resulted in the best, most sparse model (see **Supplementary Table S5**).

TABLE 1 | Linear mixed models analysis summary.

	200–250 (M1b, *df* = 13,896)			**350–450 (M4a, *df* = 13,892)**			**500–650 (M11g, *df* = 13,891)**			**SCR (M9d, *df* = 14,029)**		
Predictors	**Estimates (95% CI)**	***z***	***p***	**Estimates (95% CI)**	***z***	***p***	**Estimates (95% CI)**	***z***	***p***	**Estimates**	***z***	***p***
(Intercept) [compliment]	6.55 (5.89, 7.20)	19.58	**<0.001**	2.62 (2.03, 3.21)	8.75	**<0.001**	2.84 (2.28, 3.41)	9.79	**<0.001**	0.22 (0.19, 0.25)	13.74	**<0.001**
SpeechAct [neutral]	0.08 (−0.32, 0.47)	0.38	0.703	0.24 (−0.21, 0.69)	1.06	0.288	0.09 (−0.43, 0.61)	0.33	0.742	0.01 (0.01, 0.02)	3.97	**<0.001**
SpeechAct [insult]	0.76 (0.36, 1.16)	3.69	**<0.001**	0.65 (0.19, 1.11)	2.78	**0.005**	0.35 (−0.18, 0.88)	1.29	0.198	−0.00 (−0.01, 0.00)	-0.54	0.591
Person [3rd person]				−0.54 (−0.91, −0.17)	-2.83	**0.005**	−0.23 (−0.74, 0.27)	-0.90	0.368	0.01 (−0.00, 0.01)	1.51	0.131
BlockCent	−0.81 (−1.01, −0.60)	-7.83	**<0.001**	−0.52 (−0.87, −0.16)	-2.83	**0.005**	−0.65 (−1.04, −0.26)	-3.27	**0.001**	−0.02 (−0.04, −0.01)	-3.11	**0.002**
Block2Cent				0.50 (0.05, 0.95)	2.20	**0.028**	1.69 (1.00, 2.38)	4.80	**<0.001**	0.10 (0.08, 0.12)	12.55	**<0.001**
SpeechAct [neutral] * BlockCent				−0.52 (−1.03, −0.02)	-2.04	**0.041**	−0.54 (−1.08, 0.01)	-1.93	**0.054**			
SpeechAct [insult] * BlockCent				−0.80 (−1.31, −0.28)	-3.04	**0.002**	−1.04 (−1.60, −0.48)	-3.66	**<0.001**			
Person [3rd person] * BlockCent										0.02 (0.01, 0.03)	3.28	**0.001**
Person [3rd person] * Block2Cent							−1.24 (−2.22, −0.27)	-2.50	**0.013**	−0.06 (−0.09, −0.04)	-5.62	**<0.001**
BlockCent * Block2Cent										−0.08 (−0.10, −0.05)	-6.10	**<0.001**
	Random effects			**Random effects**			**Random effects**			**Random effects**		
σ^2	39.55			41.98			49.41			0.03		
$\tau_{00\ Participant}$	7.20			3.96			2.15			0.02		
$\tau_{00\ item}$	0.24			0.35			0.51					
ICC	0.16			0.09			0.05			0.43		
$N_{Participant}$	79			79			79			78		
N_{item}	60			60			60					
Deviance	90912.204			91702.916			93922.279			−11058.841		
Log-Likelihood	−45456.102			−45851.458			−46961.140			5529.420		

Best models for the analysis of mean ERP amplitudes in the three ERP latency ranges of interest and for the analysis of mean skin conductance response (SCR). Significant p-values are given in bold.

Individual Differences Analysis

In an additional exploratory analysis, we added the standardized scores of our various personality trait measures [the State-Trait Anxiety Inventory, the Rosenberg Self-Esteem Scale, the Interpersonal Reactivity Index, the 10-item Connor-Davidson Resilience Scale, and two Outlook and Resilience dimensions proposed by Davidson and Begley (2013)], to each best model (derived as described above) and tested whether the extended model had a better fit to the data. None of the personality trait measures significantly improved the model fit, neither in the analysis of any of the three ERP time windows nor in that of skin conductance (all p's > 0.45, see our online repository for further details). Because these are null results in exploratory analyses, we will refrain from deriving any strong inferences from them.

RESULTS

We first discuss all effects of Speech Act, for which we present mean EEG amplitude in the P2, the N400, and the LPP latency ranges, followed by the phasic skin conductance response (SCR). After this, we do the same for all Speech Act by Block effects, and then we discuss all interaction effects involving Person Described.

Speech Act

Figure 2 displays grand average ERPs, insult-compliment scalp topographies, and mean SC responses, the latter two both for the entire session as well as for the three relevant blocks (e.g., the three blocks with insults involving the participant).

Based on our earlier EEG work with morally offensive language (Van Berkum et al., 2009), we expected verbal insults to generate a larger P2, a larger N400, and a larger LPP, all relative to compliments, but presumably also relative to neutral statements. This factor was therefore included in all best final models, see **Table 1**.

P2 (200–250 ms, Anterior Region)

In line with our predictions, mean ERP amplitude in the P2 latency range was more positive to insults than to compliments ($b_{ins-com} = 0.76, z = 3.69, p_{Tukey} < 0.001$), and also more positive to insults than to neutral descriptions ($b_{ins-neu} = 0.68, z = 3.35, p_{Tukey} = 0.002$). Mean ERP amplitude in the P2 latency range for compliments and neutral descriptions did not differ ($b_{com-neu} = -0.08, z = -0.38, p_{Tukey} = 0.92$). As can be seen in **Figure 2A**, the differential insult effect, relative to compliments, has a symmetrical scalp distribution with a fronto-central maximum, consistent with other reports of emotion-induced P2 effects. We therefore interpret this as a true P2 effect.

N400 (350–450 ms, Central and Posterior Region)

Against our expectations, insults elicited a less negative mean ERP amplitude in the N400 latency range than compliments did ($b_{ins-com} = 0.65, z = 2.79, p_{Tukey} = 0.01$). Mean ERP amplitude in the N400 latency range for insults and neutral descriptions did not differ ($b_{ins-neu} = 0.41, z = 1.76, p_{Tukey} = 0.18$), nor did it for compliments vs. neutral descriptions ($b_{com-neu} = -0.24, z = -1.07, p_{Tukey} = 0.53$).

LPP (500–650 ms, Central and Posterior Region)

Also against our expectations, mean ERP amplitudes in the LPP latency range did not differ between insults and compliments ($b_{ins-com} = 0.35, z = 1.30, p_{Tukey} = 0.40$). Mean ERP amplitude in the LPP latency range for insults and neutral descriptions also did not differ ($b_{ins-neu} = 0.26, z = 0.98, p_{Tukey} = 0.59$), nor did it for compliments vs. neutral descriptions ($b_{com-neu} = -0.09, z = -0.33, p_{Tukey} = 0.94$).

Skin Conductance

Finally, mean SC response amplitudes to insults and compliments also did not differ ($b_{ins-com} = -0.00, z = -0.54, p_{Tukey} = 0.85$). Surprisingly, both of these evaluative speech acts led to a lower SCR amplitude than neutral, factually correct descriptions ($b_{ins-neu} = -0.01, z = -4.46, p_{Tukey} < 0.001$; $b_{com-neu} = -0.01, z = -3.97, p_{Tukey} < 0.001$).

Summary

In sum, across the entire session and pooled over who is described, we see insults elicit a larger P2 than compliments, but not a larger LPP, nor a larger SC response. Also, in the N400 latency range, the insult-elicited ERP-signal is more positive than the compliment-elicited ERP-signal. Note that the waveforms in **Figure 2** do suggest a more negative ERP-signal for insults than compliments around 300 ms, a timing that falls outside the scope of our predesignated latency ranges of interest. Thus, although descriptively the triphasic positive-negative-positive differential effect exhibited by the ERPs for insults vs. compliments resembles that for morally objectionable vs. acceptable statements in our earlier study (Van Berkum et al., 2009), statistical tests in the same latency ranges only confirm the early P2 effect.

Next, we examine to what extent these results are qualified by repetition of the speech acts (Block), by who is implicated (Person Described), and by the interaction between these two factors.

Speech Act as a Function of Block

In line with our critical hypothesis of a differential adaptation to repeated verbal insults and repeated compliments, we predicted that the massive repetition of these evaluative speech acts would affect at least some of the three ERP responses to insults less than it would affect the corresponding ERP responses to compliments. **Figure 3** shows the relevant model estimates (as such complementing the observed data shown in **Figure 2B**).

P2 (200–250 ms, Anterior Region)

As for mean ERP amplitude in the P2 latency range, the predicted differential adaptation to repetition was not observed. The best model included only the main effect of Speech Act and the main effect of Block, but no Speech Act by Block interaction, nor any interaction involving Person Described (see **Table 1**). As shown in **Figure 3**, mean ERP amplitudes to insults, compliments and neutral descriptions adapt over repetition in the same non-zero linear way, becoming increasingly less positive as the session progresses ($b = -0.81, z = -7.83, p_{Tukey} < 0.001$), with no significant differences between the trends. As can be seen in **Figure 2B**, and consistent with the latter, the size and scalp distribution of the differential P2-effect between insults and

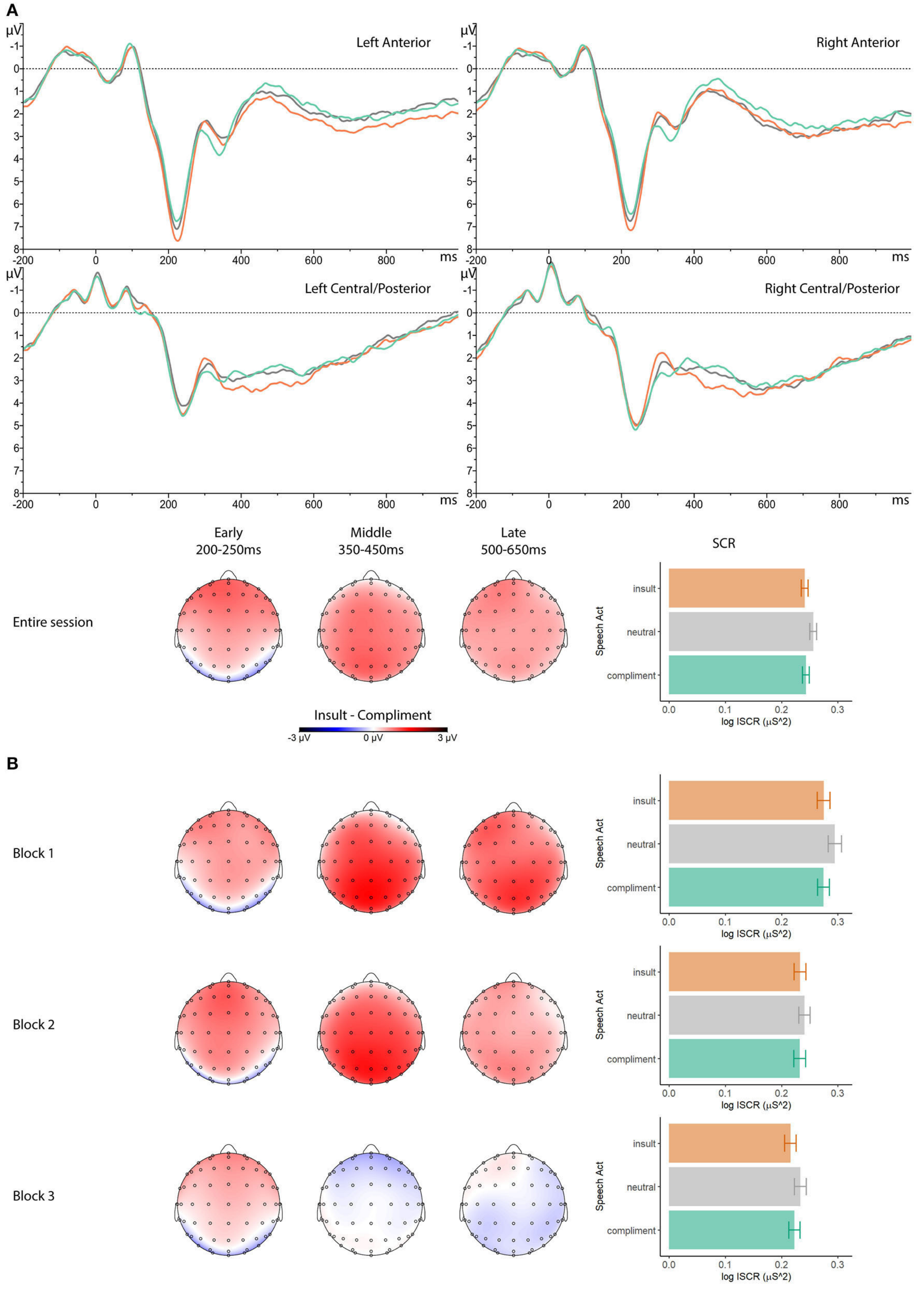

FIGURE 2 | Observed effects of speech act type across the whole session and as a function of block. **(A)** Grand average ERP waveforms (μV, negative up) in four scalp regions for insults, compliments and neutral descriptions, scalp distribution of the insults minus compliments differential ERP effect in three latency ranges of interest (μV, three left topographical maps), and the grand average event-related whole-trial skin conductance response (log $ISCR^2$, right bar graphs) for insults, compliments and neutral descriptions, all across the entire session. **(B)** Identically scaled ERP effect scalp distributions and average event-related whole-trial skin conductance response for each of the three relevant blocks in the session.

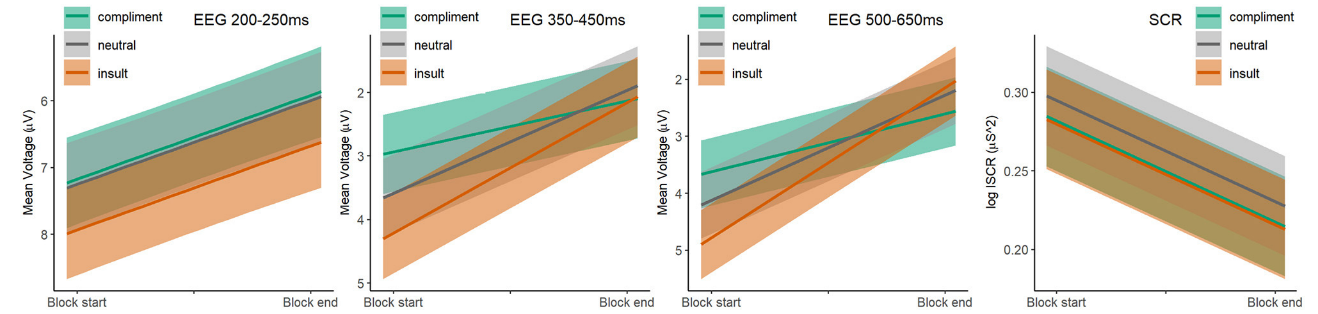

FIGURE 3 | Model estimates of Speech Act by Block interaction. Mean amplitude ERP estimated marginal means and 95% confidence intervals for the Speech Act * Block interaction for each of the three ERP latency ranges of interest (µV, three left panels, with more negative voltage up so that these panels fit with the ERP waveform panels in **Figures 2**, **4**), and the estimated marginal means and 95% confidence intervals for the Speech Act * Block interaction of the whole-trial skin conductance response (log ISCR^2, right panel). All estimates based on the final model.

compliments does not change as a function of repetition, and is clearly as visible in the last block as it is in the first one.

N400 (350–450 ms, Central and Posterior Region)

As for mean ERP amplitude in the N400 latency range, insults, compliments and neutral descriptions responded differently to repetition, albeit not in the way we had expected. The positive mean amplitudes in this latency range became less positive in a statistically linear way for all three speech act types, but with a faster rate of adaptation for insults ($b = -1.31$) than for compliments ($b = -0.52$, $b_{ins-com} = -0.80$, $z = -3.04$, $p_{Tukey} = 0.01$). As can be seen in **Figure 2B**, the widespread differential positivity for insults, compared to compliments, was visible in the first two blocks but completely absent in the last block. The rate of adaptation for insults did not statistically differ from that of neutral descriptions ($b = -1.04$; $b_{ins-neu} = -0.27$, $z = -1.06$, $p_{Tukey} = 0.54$), and the rate of adaptation for compliments also did not statistically differ from that of neutral descriptions ($b_{com-neu} = 0.52$, $z = 2.04$, $p_{Tukey} = 0.10$).

LPP (500–650 ms, Central and Posterior Region)

As for mean ERP amplitude in the LPP latency range, insults, compliments and neutral descriptions again responded differently to repetition, in a way that actually echoes responses in the N400 latency range. As in that preceding latency range, the positive mean amplitudes in the LPP latency range became less positive in a statistically linear way for all three speech act types, but with a faster rate of adaptation for insults ($b = -1.69$) than for compliments ($b = -0.65$, $b_{ins-com} = -1.04$, $z = -3.66$, $p_{Tukey} < 0.001$). As can be seen in **Figure 2B**, and consistent with the relevant slopes in **Figure 3**, the widespread differential positivity for insults, compared to compliments, was visible in the first two blocks but completely absent in the last block. The rate of adaptation for insults did not statistically differ from that of neutral descriptions ($b = -1.18$; $b_{ins-neu} = -0.51$, $z = -1.79$, $p_{Tukey} = 0.17$), and the rate of adaptation for compliments also did not statistically differ from that of neutral descriptions ($b_{com-neu} = 0.54$, $z = 1.93$, $p_{Tukey} = 0.13$).

Skin Conductance

Mean SC responses to insults, compliments and neutral descriptions all declined with repetition, at the same linear rate ($b = -0.03$, $z = -9.05$, $p_{Tukey} < 0.001$). As can be seen in **Figure 2B**, the unexpected skin conductance pattern that was observed across the entire session was also present in each of the three blocks, with the highest SC response for neutral descriptors, and no indications of differential arousal elicited by insults and compliments.

Summary

In all, our results do not provide any evidence that the massive repetition of evaluative speech acts affects responses to insults less than it affects responses to compliments. Although the P2 adapts, it does so equally for both, leaving the differential effect intact. Also, although there is differential adaptation in the N400 and LPP latency ranges, insult-induced responses adapt more than compliment-induced responses, not less. The result is that the increased ERP positivity to insults relative to compliments observed in both latency ranges has fully disappeared by the last block. The associated ERP waveforms in **Figure 2A**, the comparable scalp distributions of the insult minus compliment effect in the 350–450 ms and 500–650 ms latency ranges displayed in **Figure 2A** and, per block, in **Figure 2B**, as well as the virtually identical adaptation slopes in **Figure 3** all suggest that the most parsimonious account is one that involves the same long-lasting insult-induced positive shift, already emerging in the 350–450 ms latency range, continuing in slightly attenuated form in the 500–650 ms latency range, and vanishing for both ranges in the last block. Finally, skin conductance responses to insults, compliments, and neutral descriptors declined over blocks in a similar linear way, while retaining the unexpected elevation of arousal to neutral descriptions.

Interactions With Person Described

As revealed by the fact that none of the best statistical models retained predictors involving Person Described (see **Table 1**), the picture painted above is not significantly modulated by who the insults, compliments or neutral descriptions are about. This means that the repetition-robust P2 effect to insults relative to compliments did not depend on who was involved, and

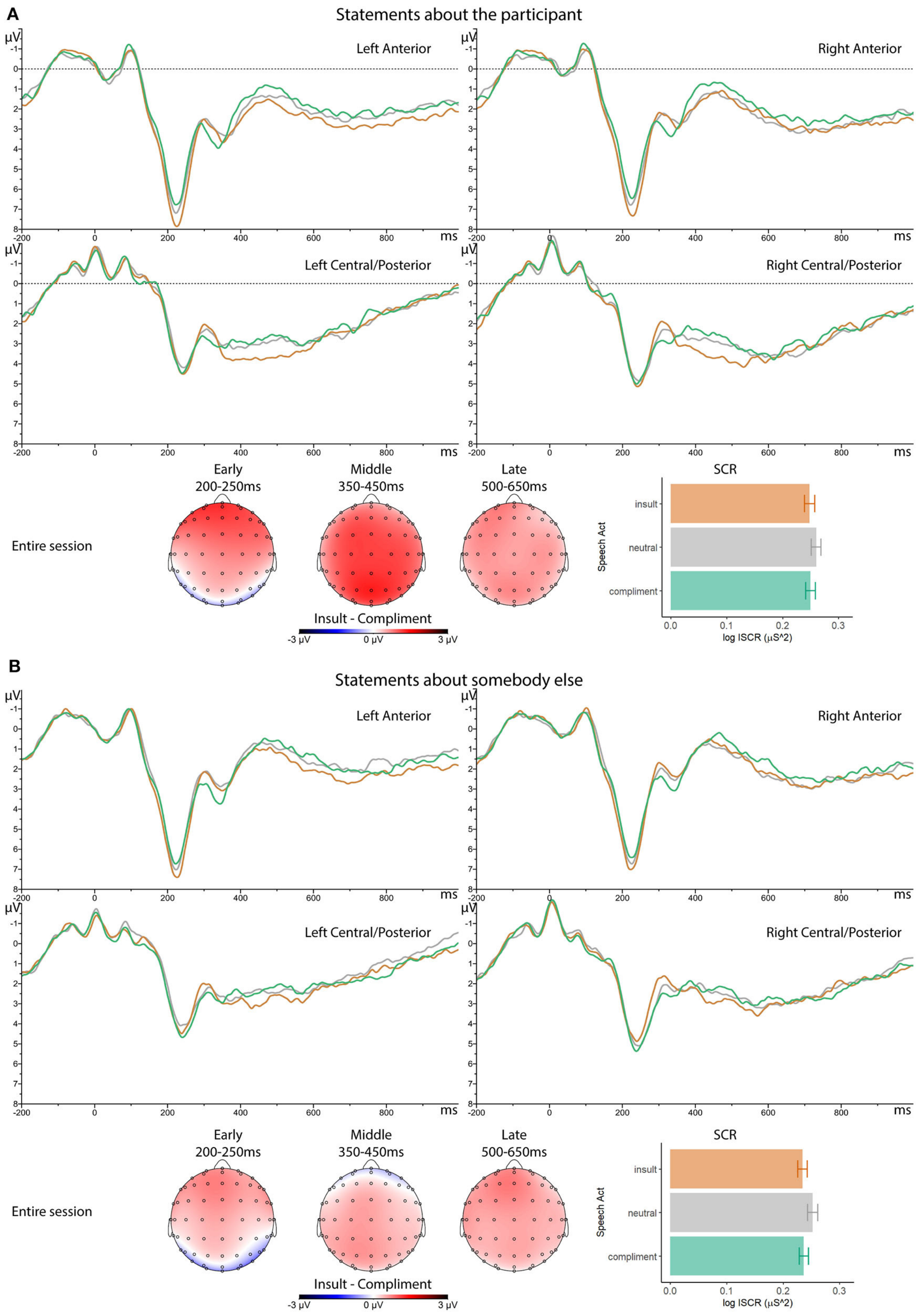

FIGURE 4 | Observed effects of speech act as a function of who is being described. Grand average ERP waveforms (μV, negative up) in four scalp regions for insults, compliments and neutral descriptions, scalp distribution of the insults minus compliments differential ERP effect in three latency ranges of interest (μV, three left topographical maps), and the grand average event-related whole-trial skin conductance response (log ISCR^2, right bar graph) for insults, compliments and neutral descriptions, all across the entire session, for **(A)** statements about the participant and **(B)** statements about somebody else.

nor did the repetition-sensitive differences between insults and compliments in later (350–450 and 500–650 ms) latency ranges. Even the unexpected higher arousal to neutral descriptors, relative to evaluative ones, did not change when statements involved the participant, rather than somebody else. The equivalence of ERP and SCR effects is illustrated in **Figure 4**. Note, in particular, the very similar triphasic positive-negative-positive response in the waveforms, and the similar P2 effect scalp distributions. Although the various insult effects seem somewhat weaker when the evaluative statements involve somebody else, this is not corroborated by the statistics.

DISCUSSION

Genuine insults violate a universal moral imperative not to inflict harm on others (Greene, 2013). Insults also threaten face or reputation, which for members of an ultrasocial species should be a highly evocative event (Goffman, 1967; Leary, 2005). We used EEG and skin conductance to explore the short-term impact of verbal insults, compared to that of verbal compliments as well as neutral factual descriptions. Partly based on our earlier ERP work on other verbal violations of moral norms (Van Berkum et al., 2009), we predicted that insults would elicit a larger P2, a larger N400, and a larger LPP than compliments (and neutral factual statements), as well as more elevated physiological arousal (see, e.g., Harris, 2004; Eilola and Havelka, 2011). We also expected that at least some of these differential insult effects would be larger when statements involved the participant, instead of somebody else. Most critically, we predicted that the EEG and skin conductance response to insults would be less affected by repetition than the equivalent responses to other speech act types.

The Early ERP Effect

The most robust insult effect in our study is the enhanced positivity emerging between 200 and 250 ms after onset of the critical insulting word. In light of prior ERP research, the most parsimonious account of this early effect is that it is a modulation of the P2 component, an anteriorly maximal short-lived ERP component that is sensitive to manipulations of attention (Luck and Hillyard, 1994; Hajcak et al., 2012). The exact functional significance and neural origin of the P2 is not yet well-understood, but research suggests that it originates early in the visual processing stream, and that an increase in amplitude indexes additional attention that is rapidly, and automatically, captured by emotional or unexpected events (Delplanque et al., 2004; Huang and Luo, 2006; Hajcak et al., 2012; Carretié, 2014). Given that insults are moral transgressions as well as verbal slaps in the face, it makes sense to find that they capture attention rapidly, at least to a larger degree than verbal compliments and neutral descriptions do.

This differential P2 effect does not change in the face of massive repetition (see **Figures 2B**, **3**). Even after two-thirds of the experiment during which readers had already seen a respectable 120 verbal insults (during which each of 10 unique insulting critical words had already been used 12 times), insults continued to generate a differential P2 effect of the same magnitude, and with the same scalp distribution. This suggests that whatever aspect of insults is responsible for capturing extra attention does so in a highly robust way. Although the stability of this differential P2 effect testifies to the fact that insults capture additional attention in a robust, repetition-insensitive way (at least within the bounds of repetition examined in this experiment), it also disconfirms our prediction that compliments would adapt sooner than insults, as the latter would predict an increasing differential effect.

Interestingly, our P2 findings suggest that the extra attention capture indexed by this effect occurs to an equal extent when insults involved the participant or somebody else (see **Figure 4**). Also in contrast to our expectations, there is no evidence that at the level of the P2, insults involving the participant adapt less to repetition than insults involving somebody else. The insensitivity of the P2 to who is being evaluated in insults and compliments suggests that the processes reflected in this early component are not influenced by the output of the compositional and referential processes that are required to compute the difference between, say, "you are a liar" and "she is a liar."

In terms of a generally accepted cascade of psycholinguistic processes (as captured, e.g., in the ALC model; Van Berkum, 2018, 2019), the most obvious explanation is that the insult-induced P2 effect has a lexical source, and simply reveals the brain's increased sensitivity to negative evaluative words, such as "liar," "disappointment," or "bitch," as the affective meaning of these signs is retrieved from long-term memory. In our study, the words used in the insults were really very negative evaluative words ("ugly," "horrible" "disappointment"), and some of them even involved taboo words (e.g., "bitch," "whore"). Such words have been observed to rapidly deliver part of that strong negative "payload"—such as by immediately capturing attention and recruiting additional processing—as part of lexical retrieval, even when presented in isolation (e.g., Wabnitz et al., 2012, 2016; see Citron, 2012, for a review). Of course, very positive words will rapidly deliver their affective meaning as part of lexical retrieval too (Citron, 2012). In our study, however, the positive evaluative words elicited a significantly smaller P2 than the negative ones, and the P2 to positive words did not differ from that elicited by critical words in factual descriptions. The P2 results in our study therefore not only reveal a stable early differential insult effect, but also testify to a negativity bias in the amount of attention that is automatically allocated to very negative vs. very positive interpersonal evaluations. Whether this should be taken to reflect valence, arousal, or a combination thereof is something we return to later.

Later ERP Effects

In earlier work (Van Berkum et al., 2009; see also Foucart et al., 2015; Hundrieser and Stahl, 2016; Peng et al., 2017), critical words in morally objectionable survey statements have elicited a larger N400 than their counterparts in morally acceptable statements. A study that compared phrasal insults to compliments (Otten et al., 2017) has revealed an N400 effect as well. In the current study, however, insulting words did not elicit a larger N400, and instead elicited a larger positivity between 350 and 450 ms. The ERP waveforms in **Figures 2** and **4** do reveal a somewhat earlier short-lived negative deflection for insulting words relative to complimenting words, around 300 ms. However, neutral factual statements descriptively elicit the same negativity, and the difference vanished before the canonical 400 ms mark.

This unexpected early negativity does not resemble the occipito-temporal EPN often reported for emotionally competent words (see Citron, 2012; Frank and Sabatinelli, 2019). The timing of this unexpected negativity falls outside our predesignated latency ranges of interest, and to not inflate error rates, we refrain from *post-hoc* statistics to corroborate it. **Figure 4** does clearly reveal that, at a descriptive level, the effect emerged in all quadrants, for speech acts involving the participant as well as somebody else. This stability bodes well for future, more hypothesis-driven examinations.

Based on our findings for morally offensive language (Van Berkum et al., 2009), and in line with other studies revealing a larger LPP for negative over positive stimuli (e.g., Schupp et al., 2004; Carretié et al., 2008; Herbert et al., 2008; Holt et al., 2009; Kissler et al., 2009; Hajcak et al., 2013; Sabatinelli et al., 2013), we had also predicted insulting words to elicit a larger LPP than compliment words (and neutral factual words). We did not obtain this insult LPP effect in our predesignated latency range of 500–650 ms. However, the waveforms shown in **Figures 2A** and **4** suggest a single insult-induced central-posterior positivity that starts around 350 ms after critical word onset and lasts until at least 600 ms. This idea is corroborated by the fact that in the 350–450 ms and 500–650 ms latency ranges, the insult effects have the same scalp distribution, as well as by the fact that mean amplitudes in the conditions involved adapt to repetition in the same way. LPP effect onsets around 300 ms are not uncommon, and even earlier LPP effects have occasionally been reported (emerging as early as 160 ms, Hajcak and Foti, 2020). The most parsimonious interpretation, therefore, is that the significant insult-induced positivity that we observe in the 350–450 ms latency range in our study is the start of an early-onsetting LPP effect[5].

The presence of an insult-induced LPP effect resonates with other work in which insults were contrasted with compliments (Otten et al., 2017), and presumably reflects rapidly increased elaborative processing of motivationally important stimuli (e.g., Cacioppo et al., 1993, 1994, 2004; Schupp et al., 2000, 2004; Smith et al., 2003; Kisley et al., 2007; Sabatinelli et al., 2007, 2013; Holt et al., 2009; Frank and Sabatinelli, 2019). However, unlike the increased emotional attention reflected in the early P2 effect, the increased elaborative processing of insults reflected in the later LPP did not withstand repetition, and had disappeared by the end of the experiment. This goes against our expectation that repeated insults might keep their sting longer than repeated compliments, and in fact reveals that at the level of the LPP, if anything, insults lose their impact on processing more rapidly than compliments. Also, this happens equally for statements involving the participant and for those involving somebody else, disconfirming our additional expectation that insults directed at the participant would adapt least of all. In fact, who was evaluated did not matter at all to LPP amplitudes in the 350–450 and 500–650 ms latency range, neither for trends over time and repetition, nor for the insult LPP effect pooled over the entire session. The hint of a stronger insult effect for insults directed at participants that is visible in **Figure 4** is not corroborated by the statistics.

Skin Conductance Effects

The skin conductance findings are puzzling. We expected that the threat to the self that is posed by insults featuring the participant's own name would generate the strongest arousal (i.e., the highest skin conductance response). In our study, however, it is the factually correct neutral descriptions that generated the strongest arousal, while insults and compliments did not differ in their arousing potential. As illustrated in **Figures 2–4** and confirmed by the statistics, this factual description effect in arousal remained the same throughout the session, and also did not differ as a function of who was involved. One explanation for the unexpected arousal to neutral factual descriptions might be that, as non-evaluative statements, they occur twice as infrequently in our experiment as the evaluative statements, and might as such perhaps require additional processing (see Citron et al., 2013, for a comparable account). However, we presented quite a few neutral statements in a single session (180), and grouped our statements in blocks of 30 of the same type, two factors that would seem to go against a simple frequentist oddball account. Furthermore, one would have to explain why these neutral oddballs noticeably increase arousal but not, for example, the P2, or the ERP amplitude in subsequent latency ranges (Polich, 2012; Hajcak and Foti, 2020).

We suspect that things are more complex than a simple frequency-based oddball account would suggest. In the context of many interpersonally evaluative statements like "Linda is horrible," "Paula is a liar," "Linda is impressive," and "Paula is an angel," non-evaluative, factually correct descriptive statements like "Linda is a person" or "Paula is a woman" may perhaps take on a meaning that makes them more complex, and more ambiguous, than we intended, and might as such have led to more arousal. Furthermore, in retrospect, some of our neutral statements, such as the 18 statements describing someone as a "person," were slightly odd—when would one ever say such a thing? Of course, this account is entirely *post-hoc*, we had not predicted our factually correct statements to generate more arousal than insults and compliments. However, we note that psycholinguistics has a long history of rediscovering that intendedly neutral stimuli are not as simple as one hoped for (see, e.g., the neutral stimuli in Rohr and Abdel Rahman, 2018). Our study may be another case in point.

Perhaps more importantly, these skin conductance results do not pattern with those in any of the ERP latency ranges that we examined in our study. This is informative as to what the insult-induced P2 and LPP effects reflect. There is an interesting debate on whether the effects of emotional words involve valence, arousal, or both (see, e.g., Bayer et al., 2010; Citron et al., 2013;

[5]In their work on morally objectionable statements, Van Berkum et al. (2009) left open the possibility that the early transgression-induced ERP positivity around 200–250 ms might also be part of an early-onsetting LPP effect, momentarily canceled out by an N400 effect. However, because the two insult-induced positivities in our study adapt in a different way, and also differ in scalp distribution, this scenario can be reasonably ruled out here.

Recio et al., 2014). In the absence of a measurable difference in physiological arousal (in an otherwise sensitive SCR measure), it is tempting to infer that our insult-induced P2 and LPP effects reflect a strong difference in valence, or some other aspect that captures the higher significance (Bradley, 2009; Hajcak and Foti, 2020) of the insults that we used.

The Bigger Picture

This brings us to an important question: what is the significance of verbal insults, presented to participants in an experiment like this? Our study took place in a lab setting where the statements were massively repeated and came from fictitious agents whose identity and fictitious opinions have no bearing on the life of the participant. This is a far cry from real life. And, of course, it had to be: research with genuinely malicious insults, in a social context in which they really make sense, and coming from people that participants really care about, is very difficult to defend. The artificial repetitiveness and grouping of speech act types, however, was a deliberate choice, because our interest was precisely in what insults would do under such repetitive conditions. Our results show that even under these highly unnatural conditions, and in the absence of a richly defined interpersonal arena, verbal insults still "get at you" and continue to do so over time, at least at some level.

But where exactly? Where in processing might the P2 and LPP insult effects observed in our study arise? The ALC model (Van Berkum, 2018, 2019; see **Figure 1**) suggests that a genuine verbal insult can be emotionally evocative at several levels in the language comprehension cascade: the social intention of the speaker (e.g., to express contempt), the referential intention (i.e., the description of you as an idiot), the stance displayed (e.g., aggressive), and the signs used (e.g., very negative and sometimes even taboo words). Because the speakers in our lab experiment are only fictitious, and because there is no real interpersonal arena (other than between experimenter and participant; Clark, 2006) it seems unlikely that our P2 and LPP effects emerge because the reader has computed a speaker's social intention or emotional stance. After all, although we hinted at a social context in our instructions, it is unlikely that our participants thought a *real* Daan or Paul was aggressively insulting people here and now. Unless an experimental situation is designed to have real interaction between real (or well-faked) interlocutors, effects at the level of social intention and speaker stance depend entirely on the imagination that participants are willing to engage in.

What about the referential intention or "situation model," the state of affairs described? Although the referential intention of an item like "Linda is horrible" is not fully defined in the lab, it may well still be an unpleasant thing to read, particularly if your name is Linda. Insulting descriptions featuring you can be experienced as slightly unpleasant even when the social intention of the speaker is *benign*, such as in teasing, playfighting, and the informal marking of intimacy amongst friends (Van Berkum, 2018)—if *that* is possible, why not in the lab? In our experiment, though, the ERP insult effects in the P2 and LPP latency range were insensitive to whether (somebody with) the participant('s name) or somebody else was being referred to in a very negative way. Our ERP effects therefore probably do not reflect the emotional consequence of a fully computed referential intention. Because a few participants later reported that the name of the other person coincided with the name of somebody they knew, we cannot exclude that any potential referential effect was somewhat diluted by this. However, when we stick to the original question, whether negative description of the self vs. somebody else matters to ERPs and skin conductance, the answer is a clear "not reliably here."[6]

Turning to the level of the signs, the verbal insults in our experiment all use very negative words (e.g., "horrible," "ugly," "bitch," "disappointment," etcetera), signs that may simply deliver their stored emotional "payload" upon retrieval and attract more emotional attention as part of that, independent of the precise sentence-level message and the social intention of the speaker on this particular occasion (Van Berkum, 2018, 2019). We already offered this as an account for the fact that insults elicited a larger P2 than compliments and continued to do so all the way through the session. On a lexical account, this robustness makes sense, because the valence or other stable aspects of emotional meaning stored with words in long-term memory is usually the result of a lifetime of experience with how words are used in one's language. Just as one cannot completely overrule the conventional non-emotional aspects of the meaning of "liar," "ugly," "bitch" or "disappointment" in a single experimental session, it might be hard to completely undo the emotional part of meaning.

Where does the emotional meaning of these words come from? As laid out in more detail elsewhere (Van Berkum, 2019), it presumably reflects experiences that pair the word with emotions triggered by other things in a sufficiently reliable way, such as other emotion-laden words that it typically collocates with (e.g., Hauser and Schwarz, 2018; Snefjella et al., 2020), the typical social intention or emotional stance of the speaker using this word, the typical impact on listeners of this word, or the nature of the things referred to by this word. Because of these correlations, the release of that stored emotional meaning as part of lexical retrieval can be seen as a quick-and-dirty context-free prediction of what the word might in the end mean in this particular context. In the confines of an artificial lab experiment, those memory-based predictions will usually be wrong. After all, in the lab, the use of, say, the word "bitch" does not predict an angry or playfighting experimenter, and "disappointment" will usually not signal social rejection—they are just words flashed on the screen as part of some bits of language to attend to, for money or course credits. Nevertheless, these memory-based predictions will probably unavoidably be made. We suspect that the insult-induced P2 observed in our study reflects the automatic emotional attention that such memory-based predictions bring about.

Just like the P2 effect, the later insult-induced LPP effect is insensitive to whether the participant or someone else was involved. However, in contrast to the former, the latter does (in

[6]*Across speech act types,* statements referring to the self did elicit larger positivities in the LPP (and preceding N400-) latency range, as well as increased skin conductance responses (i.e., arousal), than statements referring to others; see the **Supplementary Material**. This confirms the effectiveness of the self-other manipulation.

the 350–450 ms latency range) significantly diminish as insults are repeated again and again. The most probable explanation is not that the effect has a different trigger, but that it reflects the decreased allocation of more costly resources to the same lexical trigger. Apart from how it contributes to higher levels of composite and inferential meaning, a very negative word by itself elicits a cascade of different processes, some of which are automatic, and others less so. With emotionally evocative words or pictures presented in isolation, this is in fact how the LPP is construed: an index of more controlled processing (Citron, 2012), or a "more prolonged and integrated conceptual analysis" (Frank and Sabatinelli, 2019, p. 9) that is part of a larger processing cascade initiated by the same functional trigger. Some have argued that this late process involves reappraisal (Hajcak and Nieuwenhuis, 2006; Hajcak et al., 2006; Moser et al., 2006; Herbert et al., 2011), a usually effortful attempt to regulate emotional responses by reinterpreting the original stimulus. Late ERP positivities have also been associated with conscious processing more generally (see Dehaene and King, 2016), i.e., as reflecting additional "global-workspace" processing that might involve reappraisal, but can also just index some other conscious, more effortful response to the input. It is easy to imagine our participants reappraising (e.g., "this is not real, I'm in a lab"), or to otherwise reflect on the offensive materials. Either way, an insult-induced LPP effect that reflects effortful conscious processing might as such also be more vulnerable to repetition, perhaps in part because of increasing fatigue and decreasing motivation. This would explain the reliable decline of this effect in the 350–450 ms latency range.

If all this is correct, why did the very negative words in our insults not elicit a larger skin conductance response than the positive words in compliments, at least in the first part of the session, where we also observed an insult-induced LPP effect? We simply do not know. In a repetition study using emotional and neutral pictures, Schupp et al. (2006) observed a comparable dissociation between emotion-induced ERP effects that were insensitive to repetition and skin conductance effects that collapsed with repetition, and explained this in terms of the functional relevance of continued cognitive processing but decreased action preparation. However, in the face of a robust skin conductance effect to neutral, factually correct statements that lasts all the way to the end of the session, we cannot easily invoke this account for our findings.

Does the fact that insults do not increase arousal relative to the other speech acts render the insult-induced ERP-effects "non-emotional"? Note that one of the core jobs of emotion is to have *cognitive* effects, such as increased attention, stronger memorization, and changes in motivational orientation (see Van Berkum, 2022, for review). It is therefore to be expected that emotion-relevant manipulations can modulate ERP components that show up in "cognitive" psycholinguistic research as well. Furthermore, what is emotional or not is not that easy to define. In the field of emotion research, for example, surprise is actually viewed as an emotion (and, in Basic Emotions theory, even as one of the six basic emotions, Ekman and Cordaro, 2011). Also, the LPP has recently been related to the P300 (Hajcak and Foti, 2020) not because the former should be considered as irrelevant to emotion, but because both effects may index the same thing: significance to the participant. That significance may be artificially created because of task instructions, or may come naturally with the materials (Bradley, 2009). Either way, our most natural automatic response to significance—to the detection that stimuli somehow bear on our local or more general interests—is with the type of response that is the subject matter of emotion science (Van Berkum, 2022). Without claiming that the ERP components in our current study are uniquely emotional (let alone moral), we think it is safe to assume that their modulation by insulting materials has something to do with emotion.

In contrast to what we expected, though, insults in our study do not "keep their sting" across repetitions more than compliments keep their more benign emotional impact. P2 and SCR amplitudes declined at comparable rates for all speech acts, and insult-induced LPP amplitudes (spanning the 350–450 and 500–650 ms latency ranges) actually declined at a faster rate than compliment-induced ones. Also, there was no clear evidence that *who* was being insulted or complimented mattered to processing. We do not take our results to show that in real contextualized language use, people get used to insults as fast—or faster—than to compliments, nor that who is insulted does not matter. What is more likely is that, in lab experiments such as these, insults are "lexical mini-slaps" at best. Those are interesting to study, of course, and lexical effects may have a story to tell about memory traces and predictions involving emotions triggered at higher levels of real-life language use. However, in line with our ALC-model analysis, our data suggest that it takes more to study the impact of real verbal slaps in the face. In fact, perhaps more than current ethics standards allow.

Limitations

The most important limitation of our study is exactly that: our insults will have been recognized as such, but as decontextualized statements, their insulting power is limited. Interestingly, in the exit interviews, some participants did actually report feeling insulted, perhaps also because of our occasional use of swearwords. But on the whole, the absence of a real communicative situation, i.e., a real speaker addressing you about things that matter, will have placed an upper bound on the impact of our materials. Although the use of audio- or video-recorded speakers might improve things to some extent, studying real insults in the lab will remain a challenge[7]. Unless insulting statements are embedded in real (or well-faked) social interaction, processing studies are more likely to pick up on memory traces reflecting earlier experiences with real insults that participants bring to the lab than on what a particular insult can really evoke in a person, *here and now*.

A few other limitations should also be kept in mind. First, participants did not have to do anything with the statements, other than to attend to them. If the ERP effects themselves signal involvement this is not necessarily problematic, but with

[7]Presumably, decontextualization is a bigger problem for processing research on the actual emotional, "perlocutionary" effects of insults (like ours) than for 'locution-oriented' semantic-pragmatic research with sentence acceptability ratings (such as Cepollaro et al., 2021).

null results—such as no impact of who is being insulted—a taskless design can become a vulnerability. Second, we presented our stimuli in repetitive blocks. This was done on purpose, to allow us to study the response of various speech acts to massive repetition, and to test the idea that the response to insults directed at the self would be more robust than other speech act types. However, blockwise presentation will undoubtedly also have led to expectations, such as that the next item will also feature your name. Third, we have not designed our study to unconfound the effects of repetition from that of the mere passage of time. The two are always confounded in real life, and our questions were not about general adaptation effects, or the effects of mere time on task. Fourth, we assumed that the use of the participant's name would bring the statements to bear on the participant's self. Although this is not an unreasonable assumption, and echoes assumptions made in research with personal pronouns ("I," "you," etc.), it is good to keep in mind that there is some indirectness involved here. Fifth, although we successfully manipulated self- vs. other-directedness in our statements, the occasional presence of other-person names coinciding with the name of somebody they knew may have reduced our power to detect a referentially conditioned effect somewhat. Of course, the issue *was* self vs. other, so in one way, the occasional presence of names of familiar others in fact strengthens the design. At the same time, though, future studies that specifically query a referentially dependent insult effect may be more successful if they avoid the inclusion of familiar other's names. And sixth, for practical reasons that concerned the construction of our materials, we only tested female participants. We have no particular reason to believe that the observed P2 and LPP effects are gender-specific. However, some of the insulting and complimenting predicates that we used are expected to work very differently with a different gender, something to be taken into account in future replication research.

Conclusions

Genuine verbal insults violate the universal moral imperative not to harm other people. They are also threats to one's face or reputation, something that is not to be taken lightly in a species that depends on cooperation with ever changing partners. We showed that when compared to more positive evaluative statements, verbal insults elicited a very rapid P2 response that did not diminish as a function of massive repetition, and also did not depend on who was involved. The fact that verbal transgressions robustly capture attention within 250 ms after they occur, in an undefined laboratory situation where no real interaction occurs, is not only indicative of our sensitivity to undesirable social behavior, but also in line with the idea that the evaluation of such behavior is to some extent automatic. For members of an ultra-social species that depend on ever-changing cooperation, reputation, respect, and interpersonal trust, getting slapped in the face, or seeing somebody else suffer that fate, is a highly salient event, regardless of the precise context. It should be no surprise that the verbal equivalent is similarly evocative, and that the reflexive, memory-based traces of this can even be found in a psycholinguistic experiment that does not involve natural interpersonal interaction.

DATA AVAILABILITY STATEMENT

All information regarding the experimental set-up, data, and analysis of the experiment can be found in our online repository: https://doi.org/10.24416/UU01-I0U3TW.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Linguistics Chamber of the Faculty of Humanities Ethics assessment Committee at Utrecht University. The participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

Authors contributed equally to all aspects of the research, except for testing (only MS and HDM) and analysis (only MS). All authors contributed to the article and approved the submitted version.

FUNDING

Partly supported by NWO Vici grant #277-89-001 to JVB.

ACKNOWLEDGMENTS

We thank Huub van den Bergh, Hugo Quené, and Piet van Tuijl for their input during the analysis. Special thanks to Lea ter Meulen, Sanne Riemsma, and Jorik Geutjes for their help during data collection.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcomm.2022.910023/full#supplementary-material

REFERENCES

Baumeister, R. F., and Bushman, B. J. (2021). *Social Psychology and Human Nature*. Boston, MA: Cengage.

Bayer, M., Sommer, W., and Schacht, A. (2010). Reading emotional words within sentences: the impact of arousal and valence on event-related potentials. *Int. J. Psychophysiol.* 78, 299–307. doi: 10.1016/j.ijpsycho.2010.09.004

Benau, E. M., and Atchley, R. A. (2020). Some compliments (and insults) are more heartfelt. *J. Psychophysiol.* 34, 50–59. doi: 10.1027/0269-8803/a000235

Benedek, M., and Kaernbach, C. (2010). A continuous measure of phasic electrodermal activity. *J. Neurosci. Methods* 190, 80–91. doi: 10.1016/j.jneumeth.2010.04.028

Bernat, E., Bunce, S., and Shevrin, H. (2001). Event-related brain potentials differentiate positive and negative mood adjectives during both supraliminal and subliminal visual processing. *Int. J. Psychophysiol.* 42, 11–34. doi: 10.1016/S0167-8760(01)00133-7

Boucsein, W. (2012). *Electrodermal Activity*. New York, NY: Springer. doi: 10.1007/978-1-4614-1126-0

Bradley, M. M. (2009). Natural selective attention: orienting and emotion. *Psychophysiology* 46, 1–11. doi: 10.1111/j.1469-8986.2008.00702.x

Brouwer, H., Crocker, M. W., Venhuizen, N. J., and Hoeks, J. C. (2017). A neurocomputational model of the N400 and the P600 in language processing. *Cogn. Sci*. 41, 1318–1352. doi: 10.1111/cogs.12461

Brown, P., and Levinson, S. C. (1987). *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511813085

Brummelman, E., Nelemans, S. A., Thomaes, S., and Orobio de Castro, B. (2017). When parents' praise inflates, children's self-esteem deflates. *Child Dev*. 88, 1799–1809. doi: 10.1111/cdev.12936

Cacioppo, J. T., Crites, S. L., Berntson, G. G., and Coles, M. G. H. (1993). If attitudes affect how stimuli are processed, should they not affect the event-related brain potential? *Psychol. Sci*. 4, 108–112. doi: 10.1111/j.1467-9280.1993.tb00470.x

Cacioppo, J. T., Crites, S. L., Gardner, W. L., and Berntson, G. G. (1994). Bioelectrical echoes from evaluative categorizations: I. A late positive brain potential that varies as a function of trait negativity and extremity. *J. Pers. Soc. Psychol*. 67, 115–125. doi: 10.1037/0022-3514.67.1.115

Cacioppo, J. T., Larsen, J. T., Smith, N. K., and Berntson, G. G. (2004). "The affect system: what lurks below the surface of feelings?" in *Feelings and Emotions: The Amsterdam Conference*, eds A. S. R. Manstead, N. H. Frijda, and A. H. Fischer (New York, NY: Cambridge University Press), 223–242. doi: 10.1017/CBO9780511806582.014

Campbell-Sills, L., and Stein, M. B. (2007). Psychometric analysis and refinement of the Connor-Davidson Resilience Scale (CD-RISC): validation of a 10-item measure of resilience. *J. Trauma. Stress* 20, 1019–1028. doi: 10.1002/jts.20271

Carretié, L. (2014). Exogenous (automatic) attention to emotional stimuli: a review. *Cogn. Affect. Behav. Neurosci*. 14, 1228–1258. doi: 10.3758/s13415-014-0270-2

Carretié, L., Hinojosa, J. A., Albert, J., López-Martín, S., De La Gándara, B. S., Igoa, J. M., et al. (2008). Modulation of ongoing cognitive processes by emotionally intense words. *Psychophysiology* 45, 188–196. doi: 10.1111/j.1469-8986.2007.00617.x

Carretié, L., Mercado, F., Tapia, M., and Hinojosa, J. A. (2001). Emotion, attention, and the 'negativity bias', studied through event-related potentials. *Int. J. Psychophysiol*. 41, 75–85. doi: 10.1016/S0167-8760(00)00195-1

Cepollaro, B., Domaneschi, F., and Stojanovic, I. (2021). When is it ok to call someone a jerk? An experimental investigation of expressives. *Synthese* 198, 9273–9292. doi: 10.1007/s11229-020-02633-z

Citron, F. M. (2012). Neural correlates of written emotion word processing: a review of recent electrophysiological and hemodynamic neuroimaging studies. *Brain Lang*. 122, 211–226. doi: 10.1016/j.bandl.2011.12.007

Citron, F. M., Weekes, B. S., and Ferstl, E. C. (2013). Effects of valence and arousal on written word recognition: time course and ERP correlates. *Neurosci. Lett*. 533, 90–95. doi: 10.1016/j.neulet.2012.10.054

Clark, H. H. (2006). "Social actions, social commitments," in *Roots of Human Sociality: Culture, Cognition and Interaction*, eds N. J. Enfield and S. C. Levinson (Oxford: Berg), 126–150. doi: 10.4324/9781003135517-6

Codispoti, M., Bradley, M. M., and Lang, P. J. (2001). Affective reactions to briefly presented pictures. *Psychophysiology* 38, 474–478. doi: 10.1111/1469-8986.3830474

Damasio, A. (2010). *Self Comes to Mind: Constructing the Conscious Brain*. New York, NY: Pantheon/Random House.

Davidson, R. J., and Begley, S. (2013). *The Emotional Life of your Brain*. New York, NY: Hudson Street Press.

Davis, M. H. (1983). Measuring individual differences in empathy: evidence for a multidimensional approach. *J. Pers. Soc. Psychol*. 44, 113–126. doi: 10.1037/0022-3514.44.1.113

Dawson, M. E., Schell, A. M., and Filion, D. L. (2007). "The electrodermal system," in *Handbook of Psychophysiology*, eds J. T. Cacioppo, L. G. Tassinary, and G. Berntson (Cambridge University Press), 159–181.

Dehaene, S., and King, J. (2016). "Decoding the dynamics of conscious perception: the temporal generalization method," in *Micro-, Meso-and Macro-Dynamics of the Brain*, eds G. Buzsáki and Y. Christen (Cham: Springer), 85–97. doi: 10.1007/978-3-319-28802-4_7

Delplanque, S., Lavoie, M. E., Hot, P., Silvert, L., and Sequeira, H. (2004). Modulation of cognitive processing by emotional valence studied through event-related potentials in humans. *Neurosci. Lett*. 356, 1–4. doi: 10.1016/j.neulet.2003.10.014

Eilola, T. M., and Havelka, J. (2011). Behavioural and physiological responses to the emotional and taboo Stroop tasks in native and non-native speakers of English. *Int. J. Biling*. 15, 353–369. doi: 10.1177/1367006910379263

Ekman, P., and Cordaro, D. (2011). What is meant by calling emotions basic. *Emotion Review* 3, 364–370. doi: 10.1177/1754073911410740

Enfield, N. J. (2013). *Relationship Thinking: Agency, Enchrony, and Human Sociality*. New York, NY: Oxford University Press. doi: 10.1093/acprof:oso/9780199338733.001.0001

Espuny, J., Jiménez-Ortega, L., Casado, P., Fondevila, S., Muñoz, F., Hernández-Gutiérrez, D., et al. (2018). Event-related brain potential correlates of words' emotional valence irrespective of arousal and type of task. *Neurosci. Lett*. 670, 83–88. doi: 10.1016/j.neulet.2018.01.050

Fields, E. C., and Kuperberg, G. R. (2012). It's all about you: An ERP study of emotion and self-relevance in discourse. *NeuroImage*. 62, 562–574. doi: 10.1016/j.neuroimage.2012.05.003

Fields, E. C., and Kuperberg, G. R. (2015). Loving yourself more than your neighbor: ERPs reveal online effects of a self-positivity bias. *Soc. Cogn. Affect. Neurosci*. 10, 1202–1209. doi: 10.1093/scan/nsv004

Fields, E. C., and Kuperberg, G. R. (2016). Dynamic effects of self-relevance and task on the neural processing of emotional words in context. *Front. Psychol*. 6:2003. doi: 10.3389/fpsyg.2015.0200

Foti, D., Hajcak, G., and Dien, J. (2009). Differentiating neural responses to emotional pictures: evidence from temporal-spatial PCA. *Psychophysiology* 46, 521–530. doi: 10.1111/j.1469-8986.2009.00796.x

Foucart, A., Moreno, E., Martin, C. D., and Costa, A. (2015). Integration of moral values during L2 sentence processing. *Acta Psychol*. 162, 1–12. doi: 10.1016/j.actpsy.2015.09.009

Frank, D. W., and Sabatinelli, D. (2019). Hemodynamic and electrocortical reactivity to specific scene contents in emotional perception. *Psychophysiology*. 56:e13340.doi: 10.1111/psyp.13340

Frijda, N. H. (2007). *The laws of emotion*. Lawrence Erlbaum Associates Publishers.

Gantman, A. P., and Van Bavel, J. J. (2015). Moral perception. *Trends Cogn. Sci*. 19, 631–633. doi: 10.1016/j.tics.2015.08.004

Goffman, E. (1967). *Interaction ritual: Essays on Face-to-Face Interaction*. Oxford: Aldine.

Greene, J. (2013). *Moral Tribes: Emotion, Reason, and the Gap Between Us and Them*. New York, NY: Penguin Press.

Haidt, J. (2012). *The Righteous Mind: Why Good People Are Divided by Politics and Religion*. New York, NY: Pantheon/Random House.

Hajcak, G., and Foti, D. (2020). Significance? significance! empirical, methodological, and theoretical connections between the late positive potential and P300 as neural responses to stimulus significance: an integrative review. *Psychophysiology* 57, e13570. doi: 10.1111/psyp.13570

Hajcak, G., MacNamara, A., Foti, D., Ferri, J., and Keil, A. (2013). The dynamic allocation of attention to emotion: simultaneous and independent evidence from the late positive potential and steady state visual evoked potentials. *Biol. Psychol*. 92, 447–455. doi: 10.1016/j.biopsycho.2011.11.012

Hajcak, G., Moser, J. S., and Simons, R. F. (2006). Attending to affect: appraisal strategies modulate the electrocortical response to arousing pictures. *Emotion* 6, 517–522. doi: 10.1037/1528-3542.6.3.517

Hajcak, G., and Nieuwenhuis, S. (2006). Reappraisal modulates the electrocortical response to unpleasant pictures. *Cogn. Affect. Behav. Neurosci*. 6, 291–297. doi: 10.3758/CABN.6.4.291

Hajcak, G., Weinberg, A., MacNamara, A., and Foti, D. (2012). "ERPs and the study of emotion," in *The Oxford Handbook of Event-Related Potential Components*, eds S. J. Luck and E. S. Kappenman (Oxford: Oxford University Press), 441–472. doi: 10.1093/oxfordhb/9780195374148.013.0222

Halgren, E., Dhond, R., Christensen, N., van Petten, C., Marinkovic, K., Lewine, J., et al. (2002). N400-like magnetoencephalography responses modulated by semantic context, word frequency, and lexical class in sentences. *Neuroimage* 17, 1101–1116. doi: 10.1006/nimg.2002.1268

Harris, C. L. (2004). Bilingual speakers in the lab: psychophysiological measures of emotional reactivity. *J. Multiling. Multicult. Dev*. 25, 223–247. doi: 10.1080/01434630408666530

Harris, C. L., Ayçíçegí, A., and Gleason, J. B. (2003). Taboo words and reprimands elicit greater autonomic reactivity in a first language than in a second language. *Appl. Psycholinguist.* 24, 561–579. doi: 10.1017/S0142716403000286

Hauser, D. J., and Schwarz, N. (2018). How seemingly innocuous words can bias judgment: semantic prosody and impression formation. *J. Exp. Soc. Psychol.* 75, 11–18. doi: 10.1016/j.jesp.2017.10.012

Herbert, C., Junghofer, M., and Kissler, J. (2008). Event related potentials to emotional adjectives during reading. *Psychophysiology* 45, 487–498. doi: 10.1111/j.1469-8986.2007.00638.x

Herbert, C., Kissler, J., Junghöfer, M., Peyk, P., and Rockstroh, B. (2006). Processing of emotional adjectives: evidence from startle EMG and ERPs. *Psychophysiology* 43, 197–206. doi: 10.1111/j.1469-8986.2006.00385.x

Herbert, C., Pauli, P., and Herbert, B. M. (2011). Self-reference modulates the processing of emotional stimuli in the absence of explicit self-referential appraisal instructions. *Soc. Cogn. Affect. Neurosci.* 6, 653–661. doi: 10.1093/scan/nsq082

Holt, D. J., Lynn, S. K., and Kuperberg, G. R. (2009). Neurophysiological correlates of comprehending emotional meaning in context. *J. Cogn. Neurosci.* 21, 2245–2262. doi: 10.1162/jocn.2008.21151

Huang, Y., and Luo, Y. (2006). Temporal course of emotional negativity bias: an ERP study. *Neurosci. Lett.* 398, 91–96. doi: 10.1016/j.neulet.2005.12.074

Hundrieser, M., and Stahl, J. (2016). How attitude strength and information influence moral decision making: evidence from event-related potentials. *Psychophysiology* 53, 678–688. doi: 10.1111/psyp.12599

Irvine, W. B. (2013). *A Slap in the Face: How to Live in a World of Snubs, Put-downs, and Insults*. New York, NY: Oxford University Press.

Ito, T. A., Larsen, J. T., Smith, N. K., and Cacioppo, J. T. (1998). Negative information weighs more heavily on the brain: the negativity bias in evaluative categorizations. *J. Pers. Soc. Psychol.* 75, 887–900. doi: 10.1037/0022-3514.75.4.887

Jackendoff, R. (2007). A parallel architecture perspective on language processing. *Brain Res.* 1146, 2–22. doi: 10.1016/j.brainres.2006.08.111

Jay, T. (2009). The utility and ubiquity of taboo words. *Perspect. Psychol. Sci.* 4, 153–161. doi: 10.1111/j.1745-6924.2009.01115.x

Kanske, P., and Kotz, S. A. (2007). Concreteness in emotional words: ERP evidence from a hemifield study. *Brain Res.* 1148, 138–148. doi: 10.1016/j.brainres.2007.02.044

Keuleers, E., Brysbaert, M., and New, B. (2010). SUBTLEX-NL: a new measure for Dutch word frequency based on film subtitles. *Behav. Res. Methods* 42, 643–650. doi: 10.3758/BRM.42.3.643

Kisley, M. A., Wood, S., and Burrows, C. L. (2007). Looking at the sunny side of life: age-related change in an event-related potential measure of the negativity bias. *Psychol. Sci.* 18, 838–843. doi: 10.1111/j.1467-9280.2007.01988.x

Kissler, J., Assadollahi, R., and Herbert, C. (2006). Emotional and semantic networks in visual word processing: insights from ERP studies. *Prog. Brain Res.* 156, 147–183. doi: 10.1016/S0079-6123(06)56008-X

Kissler, J., Herbert, C., Winkler, I., and Junghofer, M. (2009). Emotion and attention in visual word processing—an ERP study. *Biol. Psychol.* 80, 75–83. doi: 10.1016/j.biopsycho.2008.03.004

Kunkel, A., Filik, R., Mackenzie, I. G., and Leuthold, H. (2018). Task-dependent evaluative processing of moral and emotional content during comprehension: an ERP study. *Cogn. Affect. Behav. Neurosc.* 18, 389–409. doi: 10.3758/s13415-018-0577-5

Kutas, M., and Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends Cogn. Sci.* 4, 463–470. doi: 10.1016/S1364-6613(00)01560-6

Kutas, M., and Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annu. Rev. Psychol.* 62, 621–647. doi: 10.1146/annurev.psych.093008.131123

Lang, P. J., Greenwald, M. K., Bradley, M. M., and Hamm, A. O. (1993). Looking at pictures: affective, facial, visceral, and behavioral reactions. *Psychophysiology* 30, 261–273. doi: 10.1111/j.1469-8986.1993.tb03352.x

Langeslag, S. J., and Van Strien, J. W. (2018). Early visual processing of snakes and Langeslag and Van Strien, 2018;angry faces: an ERP study. *Brain Res.* 1678, 297–303. doi: 10.1016/j.brainres.2017.10.031

Lau, E. F., Phillips, C., and Poeppel, D. (2008). A cortical network for semantics:(de) constructing the N400. *Nat. Rev. Neurosci.* 9, 920–933. doi: 10.1038/nrn2532

Lazarus, J. (2021). Negativity bias: An evolutionary hypothesis and an empirical programme. *Learn. Motivat.* 75:101731. doi: 10.1016/j.lmot.2021.101731

Leary, M. R. (2005). Sociometer theory and the pursuit of relational value: getting to the root of self-esteem. *Eur. Rev. Soc. Psychol.* 16, 75–111. doi: 10.1080/10463280540000007

Lei, Y., Dou, H., Liu, Q., Zhang, W., Zhang, Z., and Li, H. (2017). Automatic processing of emotional words in the absence of awareness: the critical role of P2. *Front. Psychol.* 8, 592. doi: 10.3389/fpsyg.2017.00592

Leuthold, H., Kunkel, A., Mackenzie, I. G., and Filik, R. (2015). Online processing of moral transgressions: ERP evidence for spontaneous evaluation. *Soc. Cogn. Affect. Neurosci.* 10, 1021–1029. doi: 10.1093/scan/nsu151

Luck, S. J., and Gaspelin, N. (2017). How to get statistically significant effects in any ERP experiment (and why you shouldn't). *Psychophysiology* 54, 146–157. doi: 10.1111/psyp.12639

Luck, S. J., and Hillyard, S. A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology* 31, 291–308. doi: 10.1111/j.1469-8986.1994.tb02218.x

MacDonald, G., and Leary, M. R. (2005). Why does social exclusion hurt? The relationship between social and physical pain. *Psychol. Bull.* 131, 202–223. doi: 10.1037/0033-2909.131.2.202

Moser, J. S., Hajcak, G., Bukay, E., and Simons, R. F. (2006). Intentional modulation of emotional responding to unpleasant pictures: an ERP study. *Psychophysiology* 43, 292–296. doi: 10.1111/j.1469-8986.2006.00402.x

Otten, M., Mann, L., van Berkum, J. J. A., and Jonas, K. J. (2017). No laughing matter: how the presence of laughing witnesses changes the perception of insults. *Soc. Neurosci.* 12, 182–193. doi: 10.1080/17470919.2016.1162194

Otten, M., and Van Berkum, J. J. A. (2007). What makes a discourse constraining? Comparing the effects of discourse message and scenario fit on the discourse-dependent N400 effect. *Brain Res.* 1153, 166–177. doi: 10.1016/j.brainres.2007.03.058

Panksepp, J., and Biven, L. (2012). *The Archaeology of Mind: Neuroevolutionary Origins of Human Emotions*. New York, NY: W. W. Norton.

Peng, X., Jiao, C., Cui, F., Chen, Q., Li, P., and Li, H. (2017). The time course of indirect moral judgment in gossip processing modulated by different agents. *Psychophysiology* 54, 1459–1471. doi: 10.1111/psyp.12893

Pessoa, L. (2013). "The impact of emotion on cognition," in *The Oxford Handbook of Cognitive Neuroscience: Volume 2: The Cutting Edges*, eds K. N. Ochsner and S. Kosslyn, (New York, NY: Oxford University Press).

Polich, J. (2012). "Neuropsychology of P300," in *The Oxford Handbook of Event-Related Potential Components*, eds S. J. Luck, and E. S. Kappenman (Oxford: Oxford University Press), 159–188.

Quené, H., and van den Bergh, H. (2004). On multi-level modeling of data from repeated measures designs: a tutorial. *Speech Commun.* 43, 103–121. doi: 10.1016/j.specom.2004.02.004

Quené, H., and van den Bergh, H. (2008). Examples of mixed-effects modeling with crossed random effects and with binomial data. *J. Mem. Lang.* 59, 413–425. doi: 10.1016/j.jml.2008.02.002

Recio, G., Conrad, M., Hansen, L. B., and Jacobs, A. M. (2014). On pleasure and thrill: the interplay between arousal and valence during visual word recognition. *Brain Lang.* 134, 34–43. doi: 10.1016/j.bandl.2014.03.009

Rohr, L., and Abdel Rahman, R. (2015). Affective responses to emotional words are boosted in communicative situations. *Neuroimage* 109, 273–282. doi: 10.1016/j.neuroimage.2015.01.031

Rohr, L., and Abdel Rahman, R. (2018). Loser! On the combined impact of emotional and person-descriptive word meanings in communicative situations. *Psychophysiology* 55, e13067. doi: 10.1111/psyp.13067

Rosenberg, M. (1965). *Society and the Adolescent Self-Image.* Princeton, NJ: Princeton University Press.

Rozin, P., and Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Pers. Soc. Psychol. Rev.* 5, 296–320. doi: 10.1207/S15327957PSPR0504_2

Sabatinelli, D., Keil, A., Frank, D. W., and Lang, P. J. (2013). Emotional perception: correspondence of early and late event-related potentials with cortical and subcortical functional MRI. *Biol. Psychol.* 92, 513–519. doi: 10.1016/j.biopsycho.2012.04.005

Sabatinelli, D., Lang, P. J., Keil, A., and Bradley, M. M. (2007). Emotional perception: correlation of functional MRI and event-related potentials. *Cereb. Cortex* 17, 1085–1091. doi: 10.1093/cercor/bhl017

Schupp, H. T., Cuthbert, B. N., Bradley, M. M., Cacioppo, J. T., Ito, T., and Lang, P. J. (2000). Affective picture processing: the late positive potential is modulated by motivational relevance. *Psychophysiology*. 37, 257–261.

Schupp, H. T., Flaisch, T., Stockburger, J., and Junghöfer, M. (2006). Emotion and attention: event-related brain potential studies. *Prog. Brain Res.* 156, 31–51. doi: 10.1016/S0079-6123(06)56002-9

Schupp, H. T., Junghöfer, M., Weike, A. I., and Hamm, A. O. (2004). The selective processing of briefly presented affective pictures: an ERP analysis. *Psychophysiology* 41, 441–449. doi: 10.1111/j.1469-8986.2004.00174.x

Schupp, H. T., and Kirmse, U. M. (2021). Case-by-case: emotional stimulus significance and the modulation of the EPN and LPP. *Psychophysiology* 58, e13766. doi: 10.1111/psyp.13766

Scott, G. G., O'Donnell, P. J., Leuthold, H., and Sereno, S. C. (2009). Early emotion word processing: evidence from event-related potentials. *Biol. Psychol.* 80, 95–104. doi: 10.1016/j.biopsycho.2008.03.010

Seasholtz, M. B., and Kowalski, B. (1993). The parsimony principle applied to multivariate calibration. *Anal. Chim. Acta* 277, 165–177. doi: 10.1016/0003-2670(93)80430-S

Shin, Y. S., and Niv, Y. (2021). Biased evaluations emerge from inferring hidden causes. *Nat. Hum. Behav.* 5, 1180–1189. doi: 10.1038/s41562-021-01065-0

Siakaluk, P. D., Pexman, P. M., Dalrymple, H. R., Stearns, J., and Owen, W. J. (2011). Some insults are more difficult to ignore: the embodied insult Stroop effect. *Lang. Cogn. Process.* 26, 1266–1294. doi: 10.1080/01690965.2010.521021

Singer, T., Seymour, B., O'Doherty, J. P., Stephan, K. E., Dolan, R. J., and Frith, C. D. (2006). Empathic neural responses are modulated by the perceived fairness of others. *Nature* 439, 466–469. doi: 10.1038/nature04271

Smith, N. K., Cacioppo, J. T., Larsen, J. T., and Chartrand, T. L. (2003). May I have your attention, please: electrocortical responses to positive and negative stimuli. *Neuropsychologia* 41, 171–183. doi: 10.1016/S0028-3932(02)00147-1

Snefjella, B., Lana, N., and Kuperman, V. (2020). How emotion is learned: semantic learning of novel words in emotional contexts. *J. Mem. Lang.* 115, 104171. doi: 10.1016/j.jml.2020.104171

Spielberger, C., Gorsuch, R. L., Lushene, R., Vagg, P., and Jacobs, G. (1983). *Manual for the State-Trait Anxiety Scale.* Palo Alto, CA: Consulting Psychologists Press.

Swaab, T. Y., Ledoux, K., Camblin, C. C., and Boudewyn, M. A. (2012). "Language related ERP components," in *Oxford Handbook of Event-Related Potential Components*, eds S. J. Luck and E. S. Kappenman (New York, NY: Oxford University Press), 397–440. doi: 10.1093/oxfordhb/9780195374148.013.0197

Tennie, C., Frith, U., and Frith, C. D. (2010). Reputation management in the age of the world-wide web. *Trends Cogn. Sci.* 14, 482–488. doi: 10.1016/j.tics.2010.07.003

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

Tomasello, M., and Herrmann, E. (2010). Ape and human cognition: what's the difference? *Curr. Direct. Psychol. Res.* 19, 3–8. doi: 10.1177/0963721409359300

Trauer, S. M., Andersen, S. K., Kotz, S. A., and Müller, M. M. (2012). Capture of lexical but not visual resources by task-irrelevant emotional words: a combined ERP and steady-state visual evoked potential study. *Neuroimage* 60, 130–138. doi: 10.1016/j.neuroimage.2011.12.016

Unkelbach, C., Koch, A., and Alves, H. (2021). Explaining negativity dominance without processing bias. *Trends Cogn. Sci.* 25, 429–430. doi: 10.1016/j.tics.2021.04.005

Van Berkum, J. J. A. (2009). "The neuropragmatics of 'simple' utterance comprehension: An ERP review" in *Semantics and Pragmatics: From Experiment to Theory* (Basingstoke: Palgrave Macmillan), 276–316.

Van Berkum, J. J. A. (2012). "The electrophysiology of discourse and conversation" in *The Cambridge Handbook of Psycholinguistics*, eds M. J. Spivey, K. McRae, and M. F. Joanisse (New York, NY: Cambridge University Press), 589–614.

Van Berkum, J. J. A. (2018). "Language comprehension, emotion, and sociality: aren't we missing something?" in *The Oxford Handbook of Psycholinguistics*, eds S. Rueschemeyer and G. Gaskell (Oxford: Oxford University Press), 644–669. doi: 10.1093/oxfordhb/9780198786825.013.28

Van Berkum, J. J. A. (2019). "Language comprehension and emotion: where are the interfaces, and who cares?" in *Oxford Handbook of Neurolinguistics*, eds G. de Zubicaray and N. O. Schiller (Oxford: Oxford University Press), 736–766. doi: 10.1093/oxfordhb/9780190672027.013.29

Van Berkum, J. J. A. (2022). "A survey of emotion theories and their relevance to language research," in *Language and Emotion: An International Handbook*, eds G. L. Schiewer, J. Altarriba, and B. C. Ng (Berlin: Mouton de Gruyter).

Van Berkum, J. J. A., Holleman, B. C., Nieuwland, M. S., Otten, M., and Murre, J. M. J. (2009). Right or wrong? The brain's fast response to morally objectionable statements. *Psychol. Sci.* 20, 1092–1099. doi: 10.1111/j.1467-9280.2009.02411.x

Vieitez, L., Haro, J., Ferré, P., Padrón, I., and Fraga, I. (2021). Unraveling the mystery about the negative valence bias: does arousal account for processing differences in unpleasant words? *Front. Psychol.* 12, 748726. doi: 10.3389/fpsyg.2021.748726

Vuilleumier, P. (2005). How brains beware: neural mechanisms of emotional attention. *Trends Cogn. Sci.* 9, 585–594. doi: 10.1016/j.tics.2005.10.011

Wabnitz, P., Martens, U., and Neuner, F. (2012). Cortical reactions to verbal abuse: event-related brain potentials reflecting the processing of socially threatening words. *Neuroreport* 23, 774–779. doi: 10.1097/WNR.0b013e328356f7a6

Wabnitz, P., Martens, U., and Neuner, F. (2016). Written threat: electrophysiological evidence for an attention bias to affective words in social anxiety disorder. *Cogn. Emot.* 30, 516–538. doi: 10.1080/02699931.2015.1019837

Wang, L., and Bastiaansen, M. (2014). Oscillatory brain dynamics associated with the automatic processing of emotion in words. *Brain Lang.* 137, 120–129. doi: 10.1016/j.bandl.2014.07.011

Weimer, N. R., Clark, S. L., and Freitas, A. L. (2019). Distinct neural responses to social and semantic violations: an N400 study. *Int. J. Psychophysiol.* 137, 72–81 B. (2020). *Statistics for Linguists: An Introduction Using R.* New York, NY; London: Routledge. doi: 10.4324/9781315165547jpsycho.2018.12.006

Winter, B. (2019). *Statistics for Linguists: An Introduction Using R.* London: Routledge.

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.

frontiers | Frontiers in Communication

TYPE Original Research
PUBLISHED 14 December 2022
DOI 10.3389/fcomm.2022.910482

Check for updates

OPEN ACCESS

EDITED BY
Francesca M. M. Citron,
Lancaster University, United Kingdom

REVIEWED BY
Bálint Forgács,
Eötvös Loránd University, Hungary
Valentina Bambini,
University School for Advanced Studies, Italy
Evelyn C. Ferstl,
University of Freiburg, Germany

*CORRESPONDENCE
Vicky Tzuyin Lai
tzuyinlai@arizona.edu

SPECIALTY SECTION
This article was submitted to Language Sciences,
a section of the journal
Frontiers in Communication

RECEIVED 01 April 2022
ACCEPTED 16 November 2022
PUBLISHED 14 December 2022

CITATION
Lai VT, van Berkum J and Hagoort P (2022) Negative affect increases reanalysis of conflicts between discourse context and world knowledge.
Front. Commun. 7:910482.
doi: 10.3389/fcomm.2022.910482

Negative affect increases reanalysis of conflicts between discourse context and world knowledge

Vicky Tzuyin Lai[1,2]*, Jos van Berkum[3] and Peter Hagoort[2]

[1]Psychology, Cognitive Science, University of Arizona, Tucson, AZ, United States, [2]Neurobiology of Language Department, Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands, [3]Languages, Literature, and Communication, Utrecht University, Utrecht, Netherlands

Introduction: Mood is a constant in our daily life and can permeate all levels of cognition. We examined whether and how mood influences the processing of discourse content that is relatively neutral and not loaded with emotion. During discourse processing, readers have to constantly strike a balance between what they know in long term memory and what the current discourse is about. Our general hypothesis is that mood states would affect this balance. We hypothesized that readers in a positive mood would rely more on default world knowledge, whereas readers in a negative mood would be more inclined to analyze the details in the current discourse.

Methods: Participants were put in a positive and a negative mood via film clips, one week apart. In each session, after mood manipulation, they were presented with sentences in discourse materials. We created sentences such as "*With the lights on you can see...*" that end with critical words (CWs) "*more*" or "*less*", where general knowledge supports "*more*", not "*less*". We then embedded each of these sentences in a wider discourse that does/does not support the CWs (a story about driving in the night vs. stargazing). EEG was recorded throughout.

Results: The results showed that first, mood manipulation was successful in that there was a significant mood difference between sessions. Second, mood did not modulate the N400 effects. Participants in both moods detected outright semantic violations and allowed world knowledge to be overridden by discourse context. Third, mood modulated the LPC (Late Positive Component) effects, distributed in the frontal region. In negative moods, the LPC was sensitive to one-level violation. That is, CWs that were supported by only world knowledge, only discourse, and neither, elicited larger frontal LPCs, in comparison to the condition where CWs were supported by both world knowledge and discourse.

Discussion: These results suggest that mood does not influence all processes involved in discourse processing. Specifically, mood does not influence lexical-semantic retrieval (N400), but it does influence elaborative processes for sensemaking (P600) during discourse processing. These results advance our understanding of the impact and time course of mood on discourse.

KEYWORDS

mood, discourse, semantics, world knowledge, ERP, N400, LPC (late positive component)

Introduction

Mood state, different from emotion, is a low-intensity, diffuse, and relatively enduring affective state (Forgas, 1995). People are in a mood as soon as they wake up and could be, for instance, cheerful, irritated, hopeful, gloomy... etc., with non-specific causes. Given the relatively enduring and long-lasting nature, people carry out daily tasks while in a certain mood. It is important to understand the effects of mood, because research has shown that mood states permeate many levels of information processing. This is the case both in obvious ways, such as prioritizing access for mood-congruent content (Egidi and Nusbaum, 2012), and also in non-obvious ways, such as loosening cognitive control to include distantly related semantic associates (Rowe et al., 2007).

Because of the high speed, incrementality, and complex interweaving of the various processes involved, much of the relevant work on mood effects in language processing has used scalp EEG (Electroencephalography)—electrical activity recorded *via* sensors on the scalp—to obtain the millisecond-by-millisecond temporal resolution needed. Similar to studies of mood on general cognition, EEG studies of mood on language have shown that mood not only affects the processing of language content but also the styles/modes of processing of readers or listeners. The present study built on this literature and used Event Related Potential (ERP) to further investigate whether and how mood influences readers' processing of discourse with language content that is relatively neutral.

Past ERP studies on mood effects on discourse focused on discourse content that is emotionally valenced, and the consensus is that mood provides affective constraint to facilitate mood-congruent content (Chung et al., 1996; Egidi and Gerrig, 2009; Egidi and Nusbaum, 2012). In Chung et al. (1996), participants were put in an optimistic or a pessimistic mood by means of personal emotional memory recall. Then, participants read stories about daily life events (e.g., a story about receiving exam grades) that ended with good and bad outcome words (*passed/failed*). They reported two ERP effects: An increased N400 (350–450 ms) for semantic- and mood- incongruent endings, and a larger LPC or Late Positive Component (300–700 ms) for mood incongruent endings. Their results indicate that participants in a pessimistic mood expected bad outcomes, and participants in an optimistic mood, good outcomes. These findings were not only replicated but also expanded in Egidi and Gerrig (2009), Egidi and Nusbaum (2012).

In terms of processing styles, past studies reported mood-specific processing styles during sentence processing (Federmeier et al., 2001; Chwilla et al., 2011; Pinheiro et al., 2013; Van Berkum et al., 2013; Wang et al., 2016). Federmeier et al. (2001), Pinheiro et al. (2013) examined mood effects on semantic categories in words in sentences. In Federmeier et al. (2001), participants were put in a positive or neutral mood. In Pinheiro et al. (2013), male participants were put in a positive, neutral, or negative mood. In both studies, participants read stories (e.g., *they wanted to make the hotel look more like a tropical resort. So, along the driveway they planted rows of...*) that ended with target words (*palms/pines/tulips*). The three target words represented three conditions: expected, within-category violation, and between-category violation. In neutral mood, they found graded N400s, largest for the between-category violation (*tulips*), intermediate for the within-category violation (*pines*), and smallest for the expected (*palms*). In positive mood, the within-category violation (*pines*) patterned with the expected (*palms*). The authors provided three possible interpretations: positive mood includes a richer set of semantic associates, positive mood flexibly accommodates unexpected/distantly related words, and positive mood entertains a plausibility-driven strategy (as opposed to a prediction-based strategy). In negative mood (Pinheiro et al., 2013), the within-category violation (*pines*) patterned with the between-category violation (*tulips*), suggesting that readers in negative mood zoom in to a narrower set of relevant semantic associates or become more critical to distantly related words.

Chwilla et al. (2011), Van Berkum et al. (2013) examined mood effects on prediction and anticipation in language. In Chwilla et al. (2011), female participants were put in a positive or negative mood, before they were presented with highly predictive sentences (e.g., *The pillows are stuffed with* ...) that continued with predicted or non-predicted critical words (*feathers/books*). In both mood states, the unpredicted (*books*) elicited a larger N400 than the predicted (*feathers*). But such N400 effect was reduced in the positive mood, compared to the negative mood. The authors suggested that positive mood allows for more prediction than negative mood does. In addition, within negative mood, there was a Late Positivity (LP) effect, larger for the unpredicted than the predicted words. The authors suggested that participants in a negative mood noticed the details and reanalyzed the unpredicted items more in this later, LP window, whereas participants in a positive mood did not. In Van Berkum et al. (2013), female participants were put in a positive or negative mood, before they read texts that contained verbs with "implicit causality biases"—that is, readers' typical expectation about who does what to whom. For example, in "*Linda apologized to David because she/he...*", readers tend to anticipate more information about Linda, which renders the pronoun "*she*" expected. However, in "*Linda praised David because he/she...*" readers tend to anticipate more information about David, which then renders the pronoun "*she*" unexpected. Such contextually unexpected pronouns have been shown to elicit larger LPs than the expected ones, and as such reveal verb-based heuristic anticipation of who will be talked about next (Van Berkum et al., 2007). Van Berkum et al. (2013) found that positive mood maintained such heuristic anticipation, whereas negative mood attenuated it. The authors speculated that a negative mood might lead the system to cut back on anticipatory referential processing of the type studied here, because the low-energy state that is typically signaled by such a mood makes such referential anticipation too resource-intensive to engage in.

The abovementioned literature supports a mood-dependent information processing style (Fredrickson, 1998) during language processing (see also Wang et al., 2016; Mills et al., 2019). Positive mood allows readers to widen semantic associates and see the bigger picture of meaning, whereas negative mood orients readers toward scrutinizing details. However, what is considered big in "big picture" may vary in language: It can stand for highly familiar, default world knowledge (e.g., knowing that more light tends to help seeing), but it can also stand for the specific discourse context that is currently configured (e.g., the astronomy context that more light tends to hinder star gazing). Relative to local processing of a word in an unfolding sentence context, *both* can in a way be considered to provide "the bigger picture" in which that processing occurs. Past non-mood studies have examined how readers juggle these two sources of knowledge, when their mood is not manipulated. Nieuwland and Van Berkum (2006) showed that all information from all sources is considered in parallel. In their study, a "local" semantic feature (animacy) in a sentence (e.g., *the peanut was salted/in love* ...) was supported or unsupported by a preceding "global" discourse context (e.g., a story about a peanut that sings and dances). They found that local semantic feature and global discourse context are processed within the same, N400 time window, suggesting that current discourse knowledge fully overrules global/default world knowledge. In contrast, Hald et al. (2007) reported that "local" discourse knowledge cannot fully override "global" world knowledge. In their study, participants read sentences that contained a critical word that was correct or incorrected based on general/global world knowledge (e.g., *The city Venice has very many canals/roundabouts* ...). These sentences were embedded in "local" discourse contexts that validate or invalidate such world knowledge (a story about this historical water city vs. a story about recent traffic control). They found a local by global interaction at the N400 time window, which indicates that while both global world knowledge and local discourse context have an effect on sentence interpretation, neither overrides the other. It appears that Nieuwland and Van Berkum (2006) viewed discourse context as being global, whereas Hald et al. (2007) viewed world knowledge as being global. An interesting question here is: which "global" or which source of knowledge would be facilitated by the "details vs. big picture" shift induced by a positive or a negative mood?

The present study examined how mood affects readers' balance between relying on world knowledge and relying on discourse knowledge. Following the abovementioned literature, we tested female participants only and manipulated their mood *via* happy and sad film clips. After mood manipulation, participants were presented with language materials. Each item contained two major pieces of world knowledge, one was cued by the discourse context and the other was cued by the critical sentence. For instance, a critical word (e.g., *more/less*) in a critical sentence (*with the lights on, you can see more/less* ...) was either supported or violated by default world knowledge cued by the critical sentential context. This critical sentence was then embedded in a discourse context that either supported the familiar world knowledge (a story about driving in the night) or supported an alternative, less familiar, but possible real world scenario (an astronomy story about stargazing). As such, our design was 2 mood (positive, negative) x 2 discourse context (supported, unsupported) x 2 critical sentence (supported, unsupported) (Table 1).

Our general predictions are that participants in a positive mood would be shifted to relying on the default world knowledge, whereas participants in a negative mood would be shifted to relying on the knowledge conveyed by discourse. As for the specific ERP components, based on the abovementioned literature, mood would impact language processing in the N400 and LPC time windows. We have mentioned these ERP components in the review above, but here we clarified the component properties and our assumptions about what they reflect. The N400 is a negative-going waveform, peaking between 200 and 600 ms, that indexes the context-dependent ease of lexical retrieval from the semantic memory (Kutas and Federmeier, 2000, 2011; Lau et al., 2008; Van Berkum, 2009; Brouwer et al., 2017). The LPC is a positive-going waveform typically occurring between 500 and 1,000 ms. The functional significance of LPC has not been settled. Some suggest that it reflects a reanalysis process of combining and recombining words for outputting sensible sentence meaning (Kuperberg, 2007). Others suggest that it reflects the demand of inference making during discourse processing (Burkhardt, 2007). Yet others associate it with an integration process that integrates all sources of information (Brouwer et al., 2012). Recently, the LPC has been linked to elaborative processes and inferences (Canal et al., 2019). Based on the synthesis of these interpretations, here we assume that LPC reflects some form of elaborative processing, e.g., more integration, or conflict resolution. Given our assumptions of these two ERP components, we expect that in the positive mood condition, words that violated default world knowledge (*with the lights on, you can see less* ...) would elicit the largest N400s, even if such reading was justified and supported by the discourse context (stargazing), following Hald et al. (2007), who used comparable materials. This expectation should also hold based on Van Berkum et al. (2013), who showed that positive mood maintains heuristics. In the negative mood condition, such discourse and sentence combination would show a reduced N400, because negative mood is more likely to pick up linguistic details in the discourse context (stargazing) to make sense of the world knowledge violation. Regarding the LPC, since both (Chwilla et al., 2011; Van Berkum et al., 2013) found that sad mood modulates LPCs (albeit the directionalities of the effects differ), we expect that readers in a negative mood would be more likely to be engaged in elaborative processing,

TABLE 1 Example stimuli.

Stimuli	CWs supported by discourse context [d+]	CWs unsupported by discourse context [d-]
CWs supported by sentence context [s+]	(1) CWs supported [d+s+] [d+]: More and more lamp posts are placed in the Netherlands. This way it is easier to see the road. This is nice for drivers. [s±] With the lights on you can see more at night.	(2) CWs partial-support [d-s+] [d-]: More and more lamp posts are placed in the Netherlands. This way it is harder to see the night sky. This is sad for astronomers. [s±]: With the lights on you can see more at night.
CWs unsupported by sentence context[s-]	(3) CWs partial-support [d+s-] [d+]: More and more lamp posts are placed in the Netherlands. This way it is harder to see the night sky. This is sad for astronomers. [s-]: With the lights on you can see less at night.	(4) CWs unsupported [d-s-] [d-]: More and more lamp posts are placed in the Netherlands. This way it is easier to see the road. This is nice for drivers. [s-] With the lights on you can see less at night.

CWs refers to critical words.

and this will be reflected in the LPCs, larger (Chwilla et al., 2011) or smaller (Van Berkum et al., 2013) LPCs.

Materials and methods

Participants

Thirty-four female, native speakers of Dutch from the Raboud University Nijmegen gave informed consent and participated in the EEG experiment for payment. Only female participants were recruited, because mood manipulation has found to be more successful in women than in men (Gross and Levenson, 1995; Federmeier et al., 2001, though see limitations in Section Limitations). Participants were assessed with the Edinburgh Inventory of Handedness (Oldfield, 1971) and the personality trait questionnaire of Positive Affect Negative Affect System (Watson and Clark, 1997). The data of several participants were excluded from the analysis, due to left-handedness ($N = 1$), PANAS personality outlier ($N = 1$), physical discomfort of illness, broken finger, and back pain ($N = 3$), technical failure ($N = 4$), and loss of trials >40% due to artifacts ($N = 1$). The remaining 24 participants (mean age = 20.4 years, range: 18–27) were right handed with normal or corrected-to-normal vision.

Design and materials

We employed a within-subject design of 2 mood (happy, sad) x 2 discourse context (support, unsupported) x 2 critical sentence (support, unsupported).

We constructed 240 quadruplets in Dutch, in the following ways (Table 1 and Supplementary material): First, we created a sentence that describes familiar world knowledge, e.g., "*with the lights on you can see more at night*". The critical word "*more*" was supported (+s) by the world knowledge. We created the condition that violates the elicited world knowledge by changing the critical word to "*less*", which was not supported (–s) by the sentence context. Next, we created a preceding discourse context whose content either reinforces the familiar knowledge ("driving in the night", d+), or goes against it ("star gazing night", d–). Thus, in condition d+s+, the critical word *more* is supported both by the familiar knowledge in the sentence (standard ideas about how light affects vision) and the discourse context (driving at night). In condition d–s–, the critical word *less* is not supported by either, as the word goes against the world knowledge (with lights on one is supposed to see better), and is also not what one would expect according to the discourse context (properly lit roads are supposed to help night driving). In condition d-s+, the critical word *more* is supported by the world knowledge, but is not what one would expect given the stargazing discourse context. It does, however, receive partial support from the sentence. Finally, in condition d+s–, although the critical word *less* is not supported by the critical sentence, it is supported by the stargazing discourse context.

We were able to recycle about a quarter of the materials from Menenti et al. (2009), Hald et al. (2007). We excluded their materials that contain scenarios that do not happen in the real world, e.g., Donald Duck, Winnie the Pooh … etc. Of the recycled ones, we edited them such that they fit our criteria described above. We also made sure to use linguistic constructions that sound natural and neutral. For example, instead of "*Amsterdam is a city that is big...*", we used "*Amsterdam is a big city...*". While both are grammatical, the former is pragmatically marked with a cleft construction (It is X that is Y), placing unnatural emphasis on the CWs.

The materials between conditions were tightly matched. In each of the 4-sentence discourse context, the first sentence introduces the topic, and is identical across all four conditions. The second and the third sentences differed between discourse types (d+) and (d–), by providing content that either supports or does not support the upcoming world knowledge cued by the critical sentence. We matched the sentence length and syntactic structure between (d+) and (d–), with minimum word differences. The critical sentential context is identical across all four conditions until the critical words (CWs hence forth), cuing world knowledge. Then, the world knowledge was either supported or not supported by the CWs, (s+) or (s–). Between (s+) and (s–), the word lengths were matched (both 7.33 letters) and the averaged log word frequencies were matched (0.85 vs. 0.84 based on CELEX (Baayen et al., 1995) and 0.80 vs. 0.80 based on SubtLex (Keuleers et al., 2010), all *p*-values n.s.). The CWs are never in a sentence-final position, nor are they also used in the discourse context.

Two pretests were conducted to verify how plausible the CWs are in the critical sentence with and without the preceding discourse contexts. Pretest 1 examined CWs in sentences without discourse contexts. The 240 sentence fragments that supported CWs and 240 sentence fragments that did not support CWs ("*With the lights on you can see more/less...* ") were divided into 2 lists *via* Latin Square rotation, such that each fragment appeared in each list only once. Within each list, the 240 items were randomized. Twenty-eight participants who did not participate in the EEG experiment or Pretest 2 (mean age 20.8, range 18–26) were randomly assigned to one of the lists, and were instructed to rate how plausible the critical word was given the preceding sentential context on a scale from 1 to 5 (1 = implausible; 5 = plausible). The mean plausibility ratings were 4.14 for (s+) and 2.38 for (s–) (Table 2). Repeated Measures ANOVA of 2 sentence x 2 list showed that list did not interact with sentence ($F < 1$), as expected. Combining lists, (s+) were more plausible than (s-) ($F(1, 239) = 77.6$, $p < 0.001$), verifying our manipulation.

Pretest 2 examined CWs in sentences with discourse contexts. The 240 [d+s+], 240 [d+s–], 240 [d–s+], and 240 [d–s–] were divided into 4 lists *via* Latin Square rotation, such that each pairing of the discourse context and the critical sentence appeared only once in each list. Forty-four participants (mean age 20.1, range 18–26) who did not take part in Pretest 1 or the main EEG experiment were randomly assigned to one of the lists each. The instructions for Pretest 2 were the same as Pretest 1. The mean plausibility ratings were 4.0 for [d+s+], 3.4 for [d–s+], 3.3 for [d+s–], and 2.2 for [d–s–] (Table 2). RM ANOVA of 2 discourse context x 2 sentence context x 4 lists showed that list did not interact with context or sentence ($F < 1$), as expected. Combining lists, There was a significant discourse context x sentence context interaction ($F(1,239) = 191.14$, $p < 0.0001$). All pairwise comparisons were significant, listed as follows. [d+s+] vs. [d+s–]: $F(1,239) = 288.91$, $p < 0.001$; [d+s–] vs. [d–s+]: $F(1,239) = 83.62$, $p < 0.001$; [d+s+] vs. [d-s-]: $F(1,239) = 118.94$, $p < 0.001$; [d+s-] vs. [d–s+]: $F(1,239) = 118.95$, $p < 0.001$; [d-s+] vs. [d-s-]: $F(1,239) = 158.73$, $p < 0.001$; [d-s+] vs. [d-s+]: $F(1,239) = 17.26$, $p < 0.001$.

Next, we divided each of the 4 lists in Pretest 2 in half into 2 sub-lists for each of the 2 mood sessions (positive, negative). That is, 120 quadruplets of sentences in each mood state. We made sure that the two sublists were comparable. The word length and frequency of the CWs between the 2 sub-lists of each list were again matched. The order of the 2 sub-lists and 2 mood sessions were counterbalanced, such that a sub-list was not always presented in one kind of mood. Then, each sub-list was divided into 5 blocks to be presented after each of the 5 mood induction video clips (more in Mood Manipulation Procedure). Within each block, the items were randomized for each participant.

To reduce session time and to avoid fatigue, we used auditory presentation of the discourse contexts that preceded the critical sentences (cf. Hald et al., 2007). One trained female Dutch speaker recorded all discourse contexts, speaking with neutral/natural intonation at a normal speaking rate. The average length of the auditory discourses is 10.5 sec (SD: 1.8 sec). The target sentences were presented visually (see Procedure for details).

We used film clips to elicit the targeted mood states, positive and negative. Meta-analyses of mood induction methods showed that films are effective in inducing the targeted emotion and that the induced emotion/mood is relatively long-lasting (Gross and Levenson, 1995; Westermann et al., 1996; Rottenberg and Gross, 2007). Based on Van Berkum et al. (2013), we used 5 film clips from a sad movie "Sophie's Choice" to induce a negative mood, and 5 film clips from a situation comedy "Friends" to induce a positive mood. Each clip lasted 3–5 min (mean 4.01 min). We verified the cheerfulness or gloominess of the film clips with a post-EEG-survey, by having EEG participants rate each film clip after the second EEG session. They were instructed to rate the films on a 1–5 scale (1 = *erg somber* "very downcast"; 5 = *erg vrolijk* "very cheerful"). The averaged film ratings were 4.5 for the "Friends" clips and were 1.6 for the "Sophie's Choice" clips (independent *t*-test: $t(30) = 16.4$, $p < 0.0001$).

Participants' mood was assessed *via* a computerized questionnaire, designed with reference to prior studies (De Vries et al., 2010; Van Berkum et al., 2013). The questionnaire contains 26 common Dutch adjectives, including 5 positive adjectives (*goed "good", tevreden "content", opgewekt "good-humored", positief "positive", vrolijk "cheerful"*), 5 negative adjectives (*down "down", slecht "bad", negatief "negative", somber "gloomy", verdrietig "sad"*), and 16 filler adjectives (*afgeleid "focused", boos "angry", geirriteerd "irritated", ongemakkelijk "uncomfortable", vermoeid "tired", zenuwachtig "nervous", slaperig "sleepy", gespannen "tense", verveeld "bored", actief "active", geconcentreerd "focused", geinteresseerd "interested", gemotiveerd "motivated", nieuwsgierig "curious", kalm "calm", ontspannen "relaxed"*). Participants were instructed to rate their mood tailored to each adjective on a 1–7 scale (1 = *Ik voelde me*

TABLE 2 Pretest results: Plausibility ratings of critical words (CWs) supported and unsupported by the critical sentence with and without discourse context.

Stimuli	No discourse	CWs supported by discourse context [d+]	CWs unsupported by discourse context [d-]
CWs supported by sentence context [s+]	4.1	4.0	3.4
CWs unsupported by sentence context [s-]	2.4	3.3	2.2

*helemaal niet "I did not at all feel"*_________; 7 = *If voelde me heel erg "I strongly feel"* _________).

Procedure

Each participant was scheduled for 2 sessions, with one week in between, at the same time-of-day and the same day-of-week. The order of mood sessions (sad first or happy first) was counterbalanced with participant number. Each session started with a 30-min EEG setup. During the setup, participants filled out the Edinburgh Inventory of Handedness and the PANAS [Positive Affect Negative Affect System, Watson et al. (1988)] personality trait questionnaire. After the setup, participants entered a soundproof, electrically shielded, and dimly lit room. They sat in a comfortable chair at a desk looking at a computer screen 70–80 cm away from their eyes. Participants were told the cover story that we were studying how concentration affects reading. They were not told that the study was about their mood states, because it is known that if participants were aware of the cause of mood change, there would be mood effects (Schwarz and Clore, 1983).

Participants first did the computerized mood rating questionnaire (baseline mood), before watching any film clip. They were asked to do the rating based on how they felt in the moment, not what they were like in general. Then, the experiment was sectioned into 5 consecutive blocks. In each block, participants watched 1 film clip, did 24 language trials, and rated their mood (in this order). Participants were instructed to watch the film clips for understanding and to listen/read the language materials attentively. Placing the mood rating at the end of each block ensured that the film-induced mood state lasted through the end of the block. The 26 adjectives on the mood questionnaire were randomized for each rating in each block, to prevent participants from memorizing their own ratings in the previous block.

In the language trials, each trial began with a discourse context presented over speakers, during which participants were told to look at the fixation sign "+" at the center of the screen. At the offset of the auditory discourse, the fixation sign remained for 1 sec, before the first word of the visual critical sentence came on the screen. The sentence was presented word-by-word, with each word presented for a length dependent duration: If a word has fewer than 8 letters, the formula was 27 ms x number of letters + 187 ms (cf. Coulson and Van Petten, 2002). If a word has more than 8 letters, the duration for 8-letter words was used. This resulted in a mean presentation duration of 370 ms for the CWs. The Inter-word Interval was a black/blank screen of 150 ms. The words were white on a black background, in Arial font, 20-point font size, and in sentence-case. The last word was presented with a period. At this point, the participant could take a tiny break or press a button to continue on to the next trial. Participants were instructed to refrain from blinking and moving during the visual presentation, but were encouraged to blink or rest their eyes between trials. There were 8 practice trials. Each EEG session lasted approximately 2 h. At the end of the 2nd session, they rated each of the film clips using a paper-and-pencil survey (cf. materials).

EEG acquisition and processing

Continuous EEG was recorded from 60 surface active electrodes placed in an elastic cap (Acticap, Brain Products, Germany) arranged in an equidistant montage (Figure 1). During recording, the left mastoid electrode served as the reference, and a forehead electrode served as the ground. A supra- to suborbital bipolar montage was used to monitor vertical eye movements (electrode 53 and VEOG), while a right to left canthal bipolar montage was used to monitor horizontal eye movements (electrodes 57 and 25). All electrode impedances were kept below 5 $K\Omega$ during recording. EEG data were amplified (0.30–100 Hz band-pass), digitized at a rate of 500 Hz with a 100 Hz high cut-off filter and a 10 second time constant.

Brain Vision Analyzer 2.0 was used to pre-process the EEG data. The EEG data were re-referenced off-line to the average of both mastoids, and low-pass filtered at 30 Hz (48 dB/oct slope). Then, the data were segmented from 200 ms before the critical word onset to 1,000 ms after, with the baseline correction from −200 to 0 ms preceding the word onset. Blinks were corrected using ICA Infomax algorithm. After that, a semi-automatic artifact rejection procedure was applied. Segments were rejected when they contained signals exceeding ±75 μV, and featured a linear drift of more than ±50 μV, beginning before the onset

 10.3389/fcomm.2022.910482

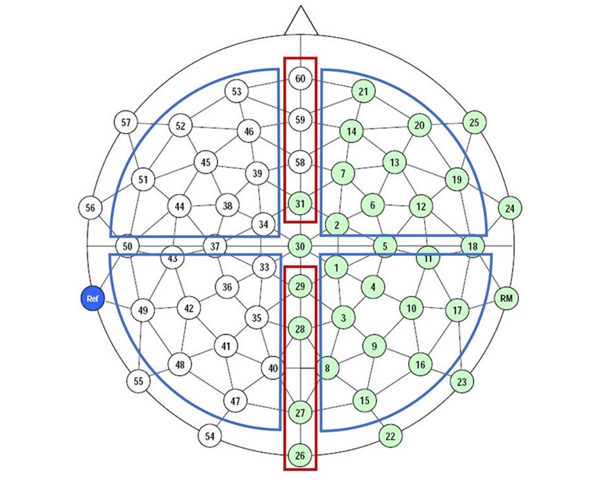

FIGURE 1
Layout of electrodes. The two red boxes indicate electrodes included in the midline analysis. The four blue boxes indicate electrodes included in the quadrant analysis.

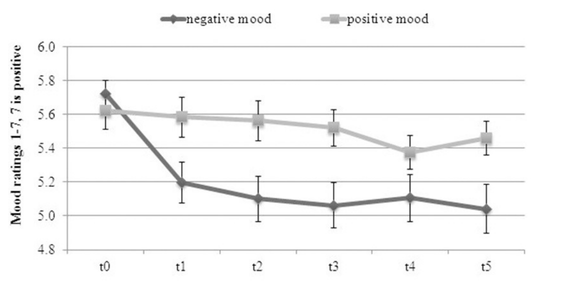

FIGURE 2
Averaged mood ratings at the baseline (t0) and the end of each block (t1, t2, t3, t4, t5), on a 1-7 scale (1 = negative mood state; 7 = positive mood state). The error bars are standard errors.

of the critical word. On average, 10% of the trials were lost. The accepted trials were averaged for each condition for each participant, and used for further statistical analysis.

Results

Mood manipulation

Mood ratings for each block were calculated by averaging the ratings from the 5 positive adjectives with transformed ratings from the 5 negative adjective. Because the scale was 1–7, we transformed the ratings by subtracting each rating from 8. In the analysis, order of mood sessions did not interact with any variable.

Figure 2 summarizes participants' mood states over time. At the baseline, there was no mood difference between the two sessions ($t(23) = 0.81$, $p = 0.426$), as expected. After watching film clips, there was significant mood difference between sessions (positive mood state vs. negative mood state in block 1: $t(23) = 2.43$, $p = 0.024$; block 2: $t(23) = 3.32$, $p = 0.003$; block 3: $t(23) = 4.75$, $p = 0.0001$; block 4: $t(23) = 2.20$, $p = 0.039$), and block 5 ($t(23) = 2.75$, $p = 0.012$). This indicates that participants were indeed in different mood states between two sessions.

Within a session, after watching the sad film clips, participants' mood dropped negative significantly relative to baseline (block 0 vs. block 1: $t(23) = 5.52$, $p < 0.0001$; block 0 vs. block 2: $t(23) = 4.93$, $p < 0.0001$; block 0 vs. block 3: $t(23) = 5.36$, $p < 0.0001$; block 0 vs. block 4: $t(23) = 4.24$, $p < 0.0001$; block 0 vs. block 5: $t(23) = 5.08$, $p < 0.0001$). However, after watching the cheerful film clips, participants' mood states were not elevated relative to baseline, but were also not down.

ERP results

The grand averages are displayed in Figure 3. Visual inspection suggested that perceptual ERP components of N1 and P2 are present, indicating normal visual processing, in both mood sessions. Following the perceptual components, there are negative-going waveforms peaking at 400 ms, identified as the N400s. The CWs unsupported by both the discourse context and the sentence context [d-s-] elicited N400s more negative than the CWs supported by both [d+s+], at the posterior sites, in both mood states. The LPCs became obvious at 600 ms and were sustained through the end of the segments at 1,000 ms. The CWs unsupported by the sentence context, the discourse context, or both ([d+s–], [d-s+], [d–s–]) elicited LPCs more positive than the CWs supported by both [d+s+], when participants were in a negative mood state (Figure 3B), but not when they were in a positive mood state (Figure 3A). These observations are supported by statistics, reported in sections 3.3 and 3.4. Topographic distributions of the N400 effects (300–500 ms) and LPC effects (600–1,000 ms) effects are displayed in Figure 4.

The mean amplitudes for the CWs from each condition in the 300–500 ms and 600–1,000 ms time windows were exported and entered into two statistical analysis: midline analysis and quadrant analysis. Midline electrodes were selected based on convention in language ERP studies. Electrodes in the quadrant regions were selected to increase coverage of the whole head. All reported numbers and p-values were Greenhouse-Geisser corrected and corrected for multiple comparisons.

N400: 300–500 ms

There is no mood modulation of N400 effects, based on the following analyses. In the midline analysis, Repeated-Measures

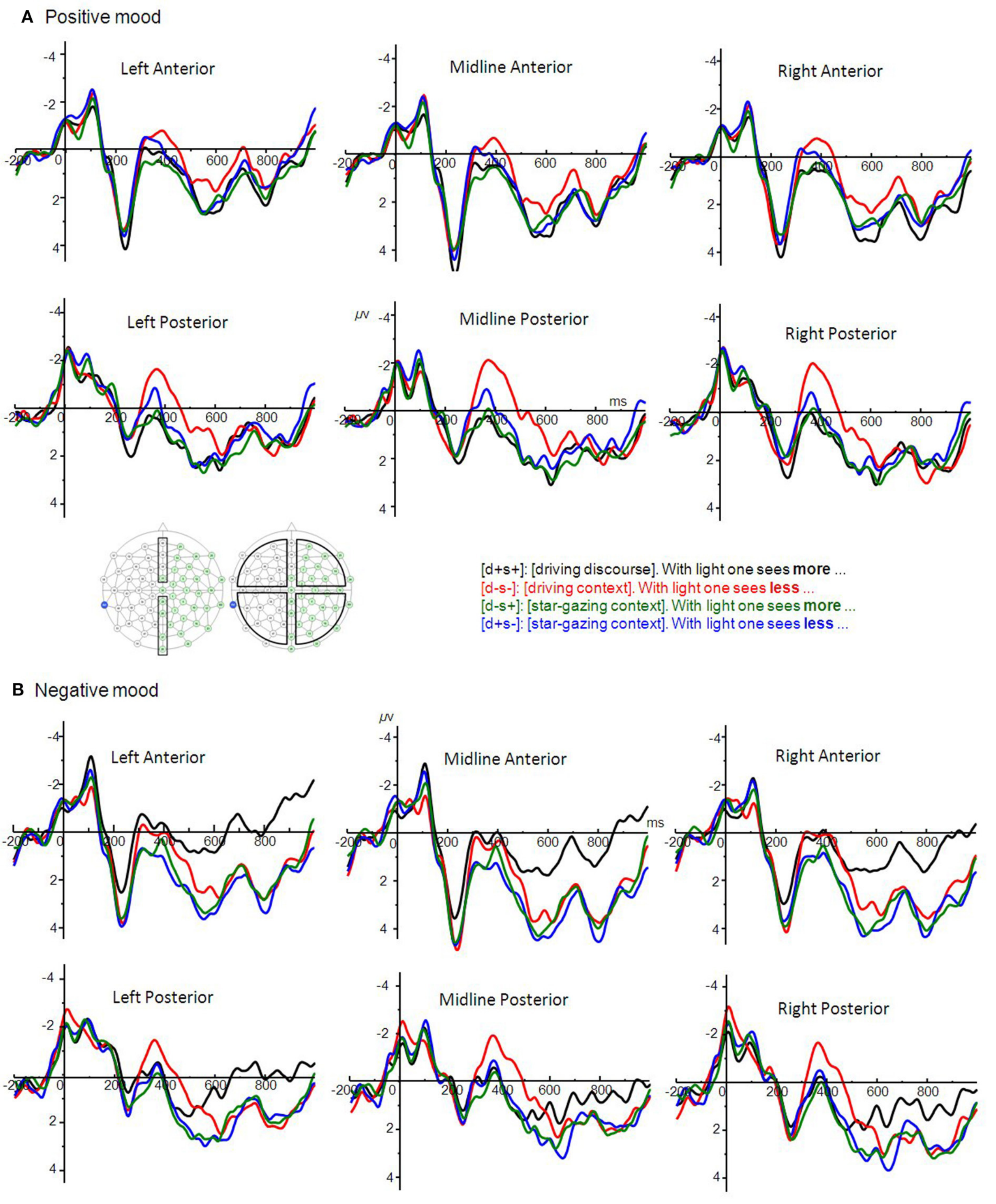

FIGURE 3
Grand ERP averages for the critical words supported by the discourse context and the sentence context ([d+s+], black line), unsupported by the discourse context and unsupported by the sentence context ([d-s-], red line), supported by the sentence context but not by the discourse context ([d-s+], green line), and supported by the discourse context but not by sentence context ([d+s-], blue line) in grouped channels in the 2 (anterior, posterior) x 3 (left, middle, right) regions, in the positive mood state **(A)** and the negative mood state **(B)**. Negative voltage is up.

10.3389/fcomm.2022.910482

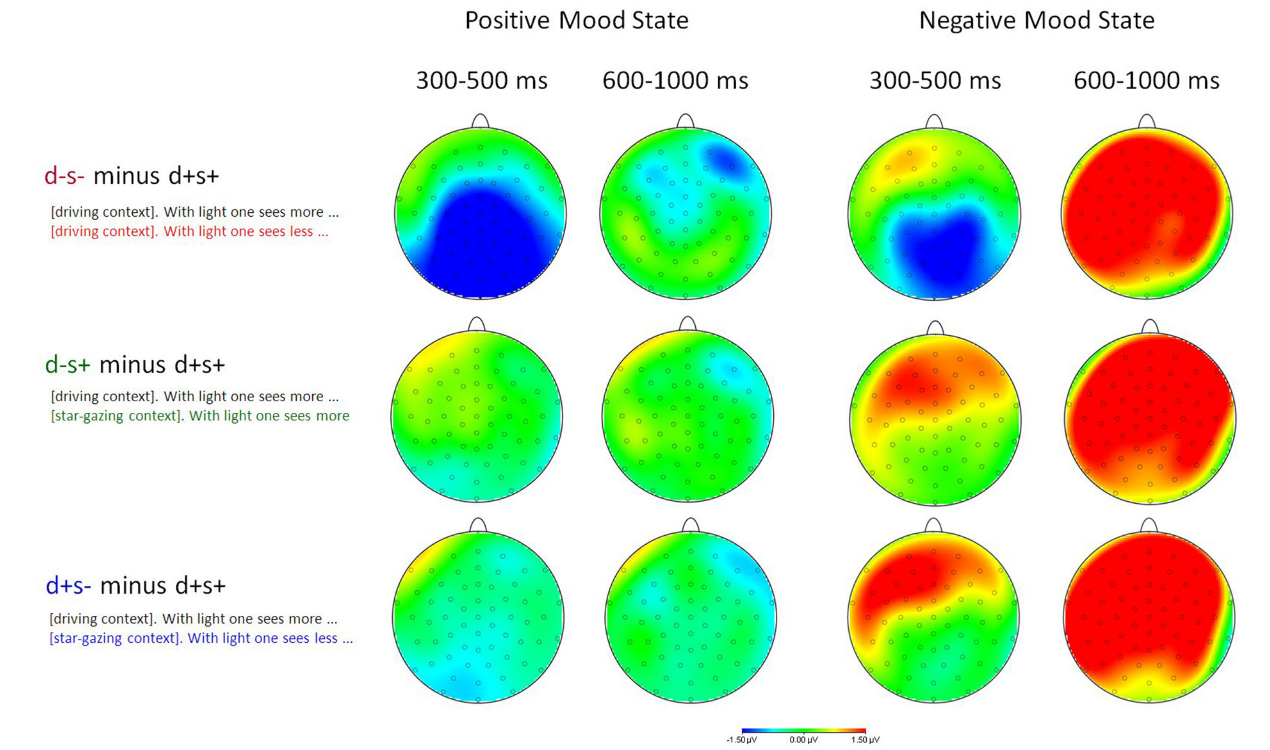

FIGURE 4
Scalp distribution of the effects obtained by subtracting the supported critical word (d+s+) from each of the other conditions (d-s-, d-s+, d+s-) in the N400 time window (300–500 ms) and the LPC (late positivity component) time window (600–1,000 ms).

(RM) ANOVAs of 2 mood (positive, negative) x 2 discourse context (supported, unsupported) x 2 sentence context (supported, unsupported) x 2 regions (frontal, posterior) x 2 order of mood revealed no 5-way interaction ($F < 1$). Combining mood, there was a significant discourse context x sentence context x region interaction ($F(1,23) = 8.74$, $p = 0.007$). Separate RM ANOVAs of 2 discourse context x 2 sentence context within each region were conducted. The N400 effects were significant in the posterior region ($F(1,23) = 6.07$, $p = 0.022$), but not in the frontal region ($F(1,23) = 0.01$, $p = 0.935$). Pairwise comparisons within the posterior region showed that the CWs unsupported by the discourse context and the sentence context [d–s–] elicited significantly larger N400s than control [d+s+] ($F(1,23) = 26.89$, $p = 0.0001$). The CWs supported by either the discourse context or the sentence context ([d+s–], [d-s+]) elicited comparable N400s to control [d+s+].

Similarly, in the quadrant analysis, RM ANOVAs of 2 mood (positive, negative) x 2 discourse context (supported, unsupported) x 2 sentence context (supported, unsupported) x 2 region_LR (left, right) x 2 region_AP (frontal, posterior) revealed no interaction at the highest level ($F(1,23) = 1.75$, $p = 0.199$). Combining mood, significant discourse context x sentence-context x region_AP interaction was observed ($F(1,23) = 10.92$, $p = 0.003$). Combining left and right, the significant discourse context x sentence context interaction came from the posterior region ($F(1,23) = 5.04$, $p = 0.035$), not from the frontal region ($F(1,23) = 0.003$, $p = 0.953$). The CWs unsupported by discourse context and sentence context [d-s-] elicited significantly more negative N400s than control [d+s+] ($F(1,23) = 18.34$, $p = 0.0001$). None of the other comparisons was significant.

Late positivity component (LPC): 600–1,000 ms

There was mood modulation of LPC effects, supported by the following statistics. In the midline analysis, RM ANOVAs of 2 mood (positive, negative) x 2 discourse context (supported, unsupported) x 2 sentence context (supported, unsupported) x 2 regions (frontal, posterior) x 2 order of mood revealed a significant 4-way interaction ($F(1,23) = 4.60$, $p = 0.043$). Breaking down the interaction, we conducted separate RM ANOVAs of 2 discourse context x 2 sentence context within each region for each mood. In the negative mood state, in the frontal region, there was a significant discourse context x sentence context interaction ($F(1,23) = 5.01$, $p = 0.035$). Pairwise comparisons showed that in the frontal region, the CWs unsupported by the discourse context [d-s+], the sentence context [d+s-], and both [d-s-] all elicited significantly more positive LPCs than control [d+s+] ([d–s+] vs. [d+s+]: $F(1,23) = 20.43$, $p = 0.0001$); ([d+s–] vs. [d+s+]: $F(1,23) = 22.56$, ($p = 0.0001$); ([d–s–] vs. [d+s+]: $F(1,23) = 17.71$, $p = 0.0001$). These effects were only marginally significant in the posterior region under the negative mood state, and were not significant in any region under the positive mood state.

Similarly, in the quadrant analysis, repeated ANOVAs of 2 mood (positive, negative) x 2 discourse context (supported, unsupported) x 2 sentence context (supported, unsupported) x 2 region_LR (left, right) x 2 region_AP (frontal, posterior) revealed a mood x discourse context x sentence context x region_AP interaction ($F(1,23) = 9.42$, $p = 0.005$). In the negative mood state, in the frontal regions, there were significant discourse context x sentence context interactions both in the left frontal region ($F(1,23) = 6.95$, p $= 0.015$) and the right frontal region ($F(1,23) = 5.55$, p $= 0.027$). Pairwise comparisons showed that the CWs unsupported by the discourse context [d-s+], the sentence context [d+s-], and both [d-s-] all elicited larger LPCs than control [d+s+] (all $p < 0.0001$). Also in the negative mood, in the posterior region, the discourse context x sentence context interaction was significant in the left posterior region ($F(1,23) = 5.40$, $p = 0.029$) and marginally significant in the right ($F(1,23) = 3.99$, $p = 0.06$). In the positive mood state, there was no LPC difference between conditions in any region.

Discussion

We conducted an ERP experiment to examine whether mood states would influence readers when they read discourse content that is not emotionally loaded. Our general hypothesis is that readers in a positive mood would rely more on default world knowledge, whereas readers in a negative mood would analyze the details in the current discourse. Female participants were put in a positive and a negative mood *via* film clips, one week apart. In each session, after mood manipulation, they were presented with vignettes that contained a critical sentence and a wider discourse context. The critical sentence contained a critical word (e.g., *more/less*) that was either supported or unsupported by the familiar world knowledge in sentential context (*with the lights on, you can see* ...). Each reading was also either supported or unsupported by the wider discourse context (a story about driving in the night/a story about stargazing).

We found that mood did not modulate the N400 effects. In both moods, CWs that were not supported by world knowledge and not supported by discourse elicited the largest N400, in comparison to the other three conditions, whose N400s were comparable to one another. Mood did modulate the LPC effects that we observed at frontal sites. In negative moods, CWs that were supported by only world knowledge, only discourse, and neither, elicited larger frontal LPCs, in comparison to the condition where CWs were supported by both world knowledge and discourse. These results partially supported our general hypothesis.

LPC (600–1,000 ms): Mood sensitive

The patterns of results in the LPC time window differed significantly between the participants' two mood sessions. Under negative mood, large and sustained LPC effects were elicited by all three experimental conditions ([d-s+], [d+s-], [d-s-]), compared to control [d+s+]. Under positive mood, there was no LPC differences between conditions. These results suggest that negative mood shifts the readers to relying more on current discourse, as opposed to relying more on default knowledge, within the LPC time window, which indexes the meaning elaboration stage (cf. Introduction). That is, readers in a negative mood are more likely to continue processing conflicted meanings from different information sources (world knowledge vs. current discourse). By processing we mean that our negatively minded readers continued to analyze and reanalyze these conflicts in an attempt to come up with a coherent output interpretation (Kuperberg, 2007), during which heavier inference drawing (Burkhardt, 2007) for elaborative processing (Canal et al., 2019) could be at work. All of these elaborative sub-processes would lead to the enhanced LPC amplitudes.

A second interesting possibility could be that the signal of conflicts in meaning triggered a "negativity bias"—i.e., the tendency to attend to negative content (Ito et al., 1998) in younger adults. Note that content-wise, our materials are actually not negatively valenced. Thus we are not suggesting negativity bias in its traditional definition. We suggest that it is the conflict between the two available information sources that might have attracted attention and invited the continued information processing in negative mood, which then led to the enhanced LPC amplitudes. If it is indeed "negativity bias" at work, then our results implicate that the definition of "negativity bias" needs to be broadened to include either (1) more attention toward (non-valenced) information as long as it is conflicting and problematic, or (2) more motivation/willingness to analyze conflicting information. The latter of the two could also become a form of rumination (Bar, 2009), fixating on the irresolvable conflicting information. Future studies will be needed to tease apart these possibilities.

Our LPC results are consistent with some but not all past ERP studies on mood on language. Our results are consistent with Chwilla et al. (2011). They found a larger LPC effect (600–800 ms) for the unpredicted CWs than the predicted CWs, in negative mood, but not in positive mood, which they suggested was due to a mood-induced reanalysis effort for the unpredicted CWs. Similar to their suggestion, we also suggest here that the negative mood nudged our participants toward a more analytical mindset. In terms of the scalp distributions of the LPC effects, ours was significant in the the frontal electrodes, whereas the LPC effect was significant in both the frontal and posterior electrodes in Chwilla et al. (2011). Such difference was likely caused by the content of the stimuli. In non-mood studies (e.g., DeLong et al., 2014), the LPC effects elicited by sentence stimuli with unpredictable but plausible CWs is more frontally distributed, whereas the LPC effects elicited by stimuli with unpredictable and anomalous CWs is distributed at posterior electrode sites. In mood studies such as ours here, we used discourse materials that described scenarios that could happen in the real world. Thus, the frontal distribution of our LPC effect makes sense. In Chwilla et al. (2011), their low predictive stimuli still had plausible endings and their LPC effect was significant at both the frontal and posterior electrodes. Synthesizing both studies, it is consistent that negative mood modulates the frontal LPCs elicited by plausible stimuli. But it is less clear what mood does for posterior LPCs elicited by implausible stimuli. This gap in knowledge is a great opportunity for future studies.

Our results might be consistent with Pinheiro et al. (2013). Pinheiro et al. (2013) did not analyze the LPC time window, likely because their study was based on Federmeier et al. (2001), who only tested positive mood and (therefore) only reported positive mood effect in the N400 time window. But Pinheiro et al. (2013) expanded the design of Federmeier et al. (2001) to include negative mood induction. In the ERPs in their negative mood [Figure 7, Pinheiro et al. (2013)], the between-category violations (*tulips*) showed a much larger LPC (600–900 ms) than their within-category violations (*pines*) in context (a tropical resort context), visually. They did not conduct analysis in this late time window. If their LPC effect was statistically significant, then their results would be consistent with ours and Chwilla et al. (2011), suggesting a more analytical processing style in negative mood. Our LPC effect (600–1,000 ms) seems less comparable to the ERP positivity effects (400–500 ms and 500–600 ms) in Van Berkum et al. (2013), which indexed anticipation heuristics and was not examined here. Overall, past and current research point to the consistent finding that readers in a negative mood tend to be more analytical of unpredicted and unexpected words.

N400 (300–500 ms): Mood insensitive

The patterns of results in the N400 time window did not differ between mood sessions. Under both moods, the [d–s–] condition (a story about driving in the night, followed by "*with the light on you see less* ...") where familiar knowledge from long term memory was not supported and without any discourse justification, elicited a larger N400 than the control [d+s+] condition (a story about driving in the night, followed by "*with the light on you see more* ..."). No N400 effect was found in the other conditions ([d+s–] and [d-s+]), both of which started with a less salient scenario (stargazing story). These results suggest that mood did not shift our readers to relying more on default world knowledge or current discourse, not in the N400 time window, which indexes context-sensitive lexical retrieval.

Combining data from both mood sessions, our N400 results only partially replicated Hald et al. (2007), where there was no mood manipulation. The main finding of Hald et al. (2007) was that neither world knowledge in long-term memory nor discourse context could completely override each other, as indexed by graded N400s. Why such discrepancy between studies? We could think of two potential explanations. The first one has to do with the differences in the materials between studies. The materials in Hald et al. (2007) consisted of a mix of fictional and real world characters and events, whereas our materials consisted of scenarios that could happen in the real world. Perhaps the authenticity of such real world knowledge attracted our participants as much as the current discourse meaning did, which then put participants' semantic system in an indeterminate state. This situation may be similar to the "Moses illusion" phenomenon, where people answer "2" to the question "how many animals of each kind did Moses take on the ark?" without noticing that it was actually Noah, not Moses, that brought animals on the ark in the original story. Notably, studies on the Moses illusion also reported a lack of N400 for a plausible semantic violation (Nieuwland and Van Berkum, 2005). A second possible explanation for the discrepancy between studies is that we used a mood manipulation, whereas Hald et al. (2007) did not. Assuming their participants were in a neutral mood, perhaps they balanced world knowledge and discourse better, not allowing one information source to override the other. And perhaps when people are in a positive or negative mood, like the participants in our study, some neural resources are occupied by the affective system, leaving insufficient resources to the cognitive system to maintain balance. These are speculations and should be tested in future studies.

Our N400 results are inconsistent with past ERP studies on mood on language. In Federmeier et al. (2001), readers in a positive mood showed a reduced N400 effect for within-category violations that had a minor difference (seeing *pines* instead of *palms* in a tropical resort context), suggesting a broader semantic activation. In Pinheiro et al. (2013), readers in a negative mood showed an increased N400 effect for the very same within-category violation, suggesting a stricter semantic activation. However, in Chwilla et al. (2011), readers in a negative mood showed reduced an N400 effect for highly unpredicted (similar to between-category violation) words in context (e.g., pillow was filled with *books* instead of *feathers*). Furthermore, a recent study (Wang et al., 2016) found that readers in a positive mood showed an enhanced N400, but only when the critical words were emphasized (focused) by context, not when they were not emphasized (non-focused). Why these discrepancies? Our current thinking post-experiment now is that at the stage of the N400 time window, mood might need to interact or work with lexical-semantic variables to make a difference: In Federmeier et al. (2001) and Pinheiro et al. (2013), the variable is the fine-grained, within-category feature. In Chwilla et al. (2011), the variable is the strong prediction for the features of the critical words. In Wang et al. (2016), the variable is focus. In our design, we did not manipulate lexical-level variables, and hence the lack of mood effects at the N400 stage.

Limitations

There are several limitations and caveats. First, we used female participants only. While this choice follows existing studies which allows us to compare our results with theirs (e.g., Chwilla et al., 2011; Wang et al., 2016), this practice limits generalization of these findings. Future studies should recruit participants from more diverse populations and mark genders in an inclusive way.

Second, while there was a significant difference between the two elicited mood states, within the positive mood session, participants' mood states were not elevated relative to baseline. It is possible that positive mood induction was not successful enough. Future studies should further examine effects of positive mood on the discourse level of language.

Third, to show mood modulation of ERP components, one might consider a correlation analysis, correlating the observed LPC effect amplitudes with mood ratings. We did not do so for two reasons: We do not have enough sample size and statistical power for a reliable correlation. In addition, the selection of electrode(s) is non-trivial. Past studies that conducted such correlation either used a carpet search approach correlating each and every electrode with mood ratings (Chwilla et al., 2011), or used only a number of electrodes that had significant amplitude results to correlate with mood ratings (Wang et al., 2016). These approaches are not ideal and could lead to incidental findings. The time window selection from anywhere from 0 to 1 second post word onset would be another issue, though recent data-driven methods might help reduce cherry picking time windows (Canal et al., 2022).

Fourth, a reviewer pointed us toward a theoretical framework, the "PET (Process, Emotion, Task) framework" (Bohn-Gettler, 2019). We did not set out to test this framework, because it was not available at the conception of this study. However, our data could certainly be related to this framework, at the situation model level under P (Process), where prior knowledge and current discourse information interact. In terms of E (Emotion), we have focused on the positive/negative valence. In terms of T (Task), we have examined constructive processing, as opposed to reproductive processing.

Finally, we used a very coarse and simplistic "valence" approach, manipulating mood and putting one in a positive or a negative mood. This probably did not capture the whole complexity surrounding the effects of mood states on information processing. Gable and Harmon-Jones (2010) encouraged researchers to also examine the motivation dimension, as they showed that positive affect that is low in approach motivational intensity (e.g., contentment) broadens cognition, whereas positive affect that is high in approach motivation (e.g., desire) narrows cognition. It would be interesting to examine the interplay between world knowledge and discourse under the influence of moods with high and low approach motivational intensity.

Conclusion

In conclusion, the current findings inform us about the effects of mood on readers' reliance on world knowledge and discourse information. Our initial predictions were that people in a positive mood would be more likely to rely on default world knowledge, whereas people in a negative mood would tend to focus on details in discourse. Our results showed that this is not entirely the case. People in a positive mood seem to entertain meaning and knowledge from both sources of real world and discourse context and are attracted to both. In contrast, people in a negative mood were shifted to relying on current discourse, reanalyzed details in all conditions that contained conflicts between different sources of information. These results advance our knowledge on the role of mood states in language meaning processing.

Data availability statement

The materials are available in the Supplementary material. Further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by Max Planck Institute for Psycholinguistics. The patients/participants provided their written informed consent to participate in this study.

Author contributions

JvB, PH, and VL co-designed the study. PH funded the study and provided the basic stimulus materials. VL constructed more stimuli and normed them and collected the EEG data and analyzed them. JvB supervised the analysis. VL prepared the manuscript with suggestions from JvB and PH. JvB and PH co-contributed to the publication fee. All authors contributed to the article and approved the submitted version.

Acknowledgments

We thank the talented and reliable research assistant Laura Arendsen for stimuli selection, construction, and revision.

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/fcomm.2022.910482/full#supplementary-material

References

Baayen, R. H., Piepenbrock, R., and Gulikers, L. (1995). *The CELEX Lexical Database (Release 2).* Distributed by the Linguistic Data Consortium, University of Pennsylvania.

Bar, M. (2009). A cognitive neuroscience hypothesis of mood and depression. *Trends Cognit. Sci.* 13, 456–463. doi: 10.1016/j.tics.2009.08.009

Bohn-Gettler, C. M. (2019). Getting a grip: The PET framework for studying how reader emotions influence comprehension. *Discourse Process.* 56, 386–401.

Brouwer, H., Crocker, M. W., Venhuizen, N. J., and Hoeks, J. C. (2017). A neurocomputational model of the N400 and the P600 in language processing. *Cognit. Sci.* 41, 1318–1352. doi: 10.1111/cogs.12461

Brouwer, H., Fitz, H., and Hoeks, J. (2012). Getting real about semantic illusions: rethinking the functional role of the P600 in language comprehension. *Brain Res.* 1446, 127–143. doi: 10.1016/j.brainres.2012.01.055

Burkhardt, P. (2007). The P600 reflects cost of new information in discourse memory. *Neuroreport* 18, 1851–1854. doi: 10.1097/WNR.0b013e3282f1a999

Canal, P., Bischetti, L., Bertini, C., Ricci, I., Lecce, S., and Bambini, V. (2022). N400 differences between physical and mental metaphors: the role of Theories of Mind. *Brain Cognit.* 161, 105879. doi: 10.1016/j.bandc.2022.105879

Canal, P., Bischetti, L., Di Paola, S., Bertini, C., Ricci, I., and Bambini, V. (2019). 'Honey, shall I change the baby?–Well done, choose another one': ERP and time-frequency correlates of humor processing. *Brain Cognit.* 132, 41–55. doi: 10.1016/j.bandc.2019.02.001

Chung, G., Tucker, D. M., West, P., Potts, G. F., Liotti, M., Luu, P., and Hartry, A. L. (1996). Emotional expectancy: brain electrical activity associated with an emotional bias in interpreting life events. *Psychophysiology* 33, 218–233. doi: 10.1111/j.1469-8986.1996.tb00419.x

Chwilla, D. J., Virgillito, D., and Vissers, C. T. W. (2011). The relationship of language and emotion: N400 support for an embodied view of language comprehension. *J. Cognit. Neurosci.* 23, 2400–2414. doi: 10.1162/jocn.2010.21578

Coulson, S., and Van Petten, C. (2002). Conceptual integration and metaphor: an event-related potential study. *Memory Cognit.* 30, 958–968. doi: 10.3758/BF03195780

De Vries, M., Holland, R. W., Chenier, T., Starr, M. J., and Winkielman, P. (2010). Happiness cools the warm glow of familiarity: psychophysiological evidence that mood modulates the familiarity-affect link. *Psychol. Sci.* 21, 321–328. doi: 10.1177/0956797609359878

DeLong, K. A., Quante, L., and Kutas, M. (2014). Predictability, plausibility, and two late ERP positivities during written sentence comprehension. *Neuropsychologia* 61, 150–162. doi: 10.1016/j.neuropsychologia.2014.06.016

Egidi, G., and Gerrig, R. J. (2009). How valence affects language processing: negativity bias and mood congruence in narrative comprehension. *Memory Cognit.* 37, 547–555. doi: 10.3758/MC.37.5.547

Egidi, G., and Nusbaum, H. C. (2012). Emotional language processing: how mood affects integration processes during discourse comprehension. *Brain lang.* 122, 199–210. doi: 10.1016/j.bandl.2011.12.008

Federmeier, K. D., Kirson, D. A., Moreno, E. M., and Kutas, M. (2001). Effects of transient, mild mood states on semantic memory organization and use: an event-related potential investigation in humans. *Neurosci. Lett.* 305, 149–152. doi: 10.1016/S0304-3940(01)01843-2

Forgas, J. P. (1995). Mood and judgment: the affect infusion model (AIM). *Psychol. Bull.* 117, 39. doi: 10.1037/0033-2909.117.1.39

Fredrickson, B. L. (1998). What good are positive emotions?. *Rev. Gener. Psychol.* 2, 300–319. doi: 10.1037/1089-2680.2.3.300

Gable, P., and Harmon-Jones, E. (2010). The motivational dimensional model of affect: implications for breadth of attention, memory, and cognitive categorisation. *Cognit. Emot.* 24, 322–337. doi: 10.1080/02699930903378305

Gross, J. J., and Levenson, R. W. (1995). Emotion elicitation using films. *Cognit. Emot* 9, 87–108. doi: 10.1080/02699939508408966

Hald, L. A., Steenbeek-Planting, E. G., and Hagoort, P. (2007). The interaction of discourse context and world knowledge in online sentence comprehension. Evidence from the N400. *Brain Res.* 1146, 210–218. doi: 10.1016/j.brainres.2007.02.054

Ito, T. A., Larsen, J. T., Smith, N. K., and Cacioppo, J. T. (1998). Negative information weighs more heavily on the brain: the negativity bias in evaluative categorizations. *J. Personal. Soc. Psychol.* 75, 887. doi: 10.1037/0022-3514.75.4.887

Keuleers, E., Brysbaert, M., and New, B. (2010). SUBTLEX-NL: a new measure for Dutch word frequency based on film subtitles. *Behav. Res. Methods.* 42, 643–650. doi: 10.3758/BRM.42.3.643

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: challenges to syntax. *Brain Res.* 1146, 23–49. doi: 10.1016/j.brainres.2006.12.063

Kutas, M., and Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends Cognit. Sci.* 4, 463– 470. doi: 10.1016/S1364-6613(00)01560-6

Kutas, M., and Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Ann. Rev. Psychol.* 62, 621–647. doi: 10.1146/annurev.psych.093008.131123

Lau, E. F., Phillips, C., and Poeppel, D. (2008). A cortical network for semantics:(de) constructing the N400. *Nat. Rev. Neurosci.* 9, 920–933. doi: 10.1038/nrn2532

Menenti, L., Petersson, K. M., Scheeringa, R., and Hagoort, P. (2009). When elephants fly: differential sensitivity of right and left inferior frontal gyri to discourse and world knowledge. *J. Cognit. Neurosci.* 21, 2358–2368. doi: 10.1162/jocn.2008.21163

Mills, C., Wu, J., and D'Mello, S. (2019). Being sad is not always bad: the influence of affect on expository text comprehension. *Disc. Process.* 56, 99–116. doi: 10.1080/0163853X.2017.1381059

Nieuwland, M. S., and Van Berkum, J. J. (2005). Testing the limits of the semantic illusion phenomenon: ERPs reveal temporary semantic change deafness in discourse comprehension. *Cognit. Brain Res.* 24, 691–701. doi: 10.1016/j.cogbrainres.2005.04.003

Nieuwland, M. S., and Van Berkum, J. J. (2006). When peanuts fall in love: N400 evidence for the power of discourse. *J. Cognit. Neurosci.* 18, 1098–1111. doi: 10.1162/jocn.2006.18.7.1098

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4

Pinheiro, A. P., del Re, E., Nestor, P. G., McCarley, R. W., Gonçalves, O. F., and Niznikiewicz, M. (2013). Interactions between mood and the structure of semantic memory: event-related potentials evidence. *Soc. Cognit. Affect. Neurosci.* 8, 579–594. doi: 10.1093/scan/nss035

Rottenberg, J., and Gross, J. J. (2007). Emotion and emotion regulation: a map for psychotherapy researchers. doi: 10.1111/j.1468-2850.2007.00093.x

Rowe, G., Hirsh, J. B., and Anderson, A. K. (2007). Positive affect increases the breadth of attentional selection. *Proc. Natl. Acad. Sci.* 104, 383–388. doi: 10.1073/pnas.0605198104

Schwarz, N., and Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: informative and directive functions of affective states. *J. Personal. Soc. Psychol.* 45, 513–523. doi: 10.1037/0022-3514.45.3.513

Van Berkum, J. J., Koornneef, A. W., Otten, M., and Nieuwland, M. S. (2007). Establishing reference in language comprehension: an electrophysiological perspective. *Brain Res.* 1146, 158–171. doi: 10.1016/j.brainres.2006.06.091

Van Berkum, J. J., De Goede, D., Van Alphen, P., Mulder, E., and Kerstholt, J. H. (2013). How robust is the language architecture? The case of mood. *Front. Psychol.* 4, 505. doi: 10.3389/fpsyg.2013.00505

Van Berkum, J. J. A. (2009). "The neuropragmatics of 'simple' utterance comprehension: an ERP review," in *Semantics and Pragmatics: From Experiment to Theory* (London: Palgrave Macmillan), pp. 276–316.

Wang, L., Zhou, B., Zhou, W., and Yang, Y. (2016). Odor-induced mood state modulates language comprehension by affecting processing strategies. *Sci. Rep.* 6, 1–13. doi: 10.1038/srep36229

Watson, D., and Clark, L. A. (1997). Measurement and mismeasurement of mood: Recurrent and emergent issues. J Pers Assess, 68, 267–296. doi: 10.1207/s15327752jpa6802_4

Watson, D., Clark, L. A., and Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *J. Personal. Soc. Psychol.* 54, 1063. doi: 10.1037/0022-3514.54.6.1063

Westermann, R., Spies, K., Stahl, G., and Hesse, F. W. (1996). Relative effectiveness and validity of mood induction procedures: a meta-analysis. *Eur. J. Soc. Psychol.* 26, 557–580. doi: 10.1002/(SICI)1099-0992(199607)26:4<557::AID-EJSP769>3.0.CO;2-4

Frontiers in Communication

Investigates the power of communication across culture and society

A cross-disciplinary journal that advances our understanding of the global communication revolution and its relevance across social, economic and cultural spheres.

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