

# SUB- AND SUPRA-SECOND TIMING: BRAIN, LEARNING AND DEVELOPMENT

EDITED BY : Lihan Chen, Yan Bao and Marc Wittmann  
PUBLISHED IN : Frontiers in Psychology







# frontiers

## Frontiers Copyright Statement

© Copyright 2007-2016 Frontiers Media SA. All rights reserved.

All content included on this site, such as text, graphics, logos, button icons, images, video/audio clips, downloads, data compilations and software, is the property of or is licensed to Frontiers Media SA ("Frontiers") or its licensees and/or subcontractors. The copyright in the text of individual articles is the property of their respective authors, subject to a license granted to Frontiers.

The compilation of articles constituting this e-book, wherever published, as well as the compilation of all other content on this site, is the exclusive property of Frontiers. For the conditions for downloading and copying of e-books from Frontiers' website, please see the Terms for Website Use. If purchasing Frontiers e-books from other websites or sources, the conditions of the website concerned apply.

Images and graphics not forming part of user-contributed materials may not be downloaded or copied without permission.

Individual articles may be downloaded and reproduced in accordance with the principles of the CC-BY licence subject to any copyright or other notices. They may not be re-sold as an e-book.

As author or other contributor you grant a CC-BY licence to others to reproduce your articles, including any graphics and third-party materials supplied by you, in accordance with the Conditions for Website Use and subject to any copyright notices which you include in connection with your articles and materials.

All copyright, and all rights therein, are protected by national and international copyright laws.

The above represents a summary only. For the full conditions see the Conditions for Authors and the Conditions for Website Use.

ISSN 1664-8714

ISBN 978-2-88919-898-6

DOI 10.3389/978-2-88919-898-6

## 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: [researchtopics@frontiersin.org](mailto:researchtopics@frontiersin.org)

# SUB- AND SUPRA-SECOND TIMING: BRAIN, LEARNING AND DEVELOPMENT

Topic Editors:

**Lihan Chen**, Peking University, China

**Yan Bao**, Peking University, China

**Marc Wittmann**, Institute for Frontier Areas of Psychology and Mental Health, Germany



The cover figure tries to depict the essence of time perception as ‘persistent change’ and experience of ‘flow’, typified by waves in the human face and changes from dark colors to light ones in the background.

Figure by Teng Zhang, inspired by Salvador Dali <Persistent of Memory>.

as compared to when humans only passively perceive the same perceptual scenario (Bao et al., 2015; Chen & Vroomen, 2013).

Time perception in the range of milliseconds to a few seconds is essential for many important sensory and perceptual tasks including speech perception, motion perception, motor coordination, and cross-modal interaction. For the brain to be in synchrony with the environment, the physical differences in the speeds of light and sound, as well as stimuli from other modalities such as odors, must be processed and coordinated (Pöppel & Bao 2014; Bao et al., 2015).

Time is a subjective feeling that is modulated by emotional states which trigger temporal distortions (temporal dilation vs. contraction) (Wittmann et al., 2014), hence give rise to subjective time that may be different to event time as initially registered in the brain. Recent research suggests that time perception in a multisensory world is subject to prior task experience and shaped by (statistical) learning processes. Humans are active learners. That is, the engagement of the own body in a timing task within a perceptual-action loop will make a noticeable difference in timing performance,

This Research Topic of “Sub-and supra-second timing: brain, learning and development” has integrated sixteen submissions of novel research on sub- and supra-timing. We have categorized the papers in this topic into the following four themes, from which we can deduce trends of research about multisensory timing in the sub- and supra-second range.

## Sensory Timing, Interaction and Reliability

A central debate in sensory timing is whether it is subserved by a centralized timing mechanism or distinctive/modular processing (Ivry & Schlerf, 2008). We included five papers underlying this theme. Di Luca investigated how judgments of perceived duration are influenced by the properties of the signals that define the intervals. They found that timing distortion is attributed to both intervals (isochronous vs anisochronous) and filling types (empty vs. filled) (Horr & Di Luca, 2015). Cai and Eagleman asked themselves how the brain forms a representation of duration when each of two simultaneously presented stimuli influence perceived duration in different ways. They attributed the perceived averaged duration of simultaneously occurring visual stimuli to the weightings of the elementary (individual) stimuli, although the weighting performance did not fully predict statistically optimal integration (Cai & Eagleman, 2015). Birngruber, Schröter and Ulrich examined the effects of stimulus repetition vs. stimulus novelty on perceived duration. They substantiated the view that changes of simple, that is, semantically meaningless stimuli lead to shorter perceived duration of repeated as compared to novel stimuli (Birngruber, Schroter, & Ulrich, 2015). In the conceptual framework of the distinct timing hypothesis, Rammsayer et al. showed a gradual transition from a purely modality-specific, sensory-automatic to a more cognitive, amodal timing mechanism, by viewing the evidence that the prevalence of precision of auditory over visual timing disappeared when the temporal range is controlled (Rammsayer, Bortner, & Troche, 2015). Yue et al. explored the effects of olfactory events upon reproduced time durations in auditory and visual modalities, and found that the biased timing in target stimuli (auditory and visual) could be accounted for by a framework of attentional deployment between the inducers (odors) and emotionally neutral stimuli (visual dots and sound beeps) (Yue, Gao, Chen, & Wu, 2016). However, the mechanisms of distinct timing vs. centralized timing are to be determined.

## Adaptive Representation of Time, Learning and Temporal Prediction

Golan and Zakay probed the duality of temporal encoding –the intrinsic and extrinsic representation of time- using fMRI. They exposed participants to stimuli with different temporal variance and found neural activation (within category-selective brain regions) increase as a function of increase in temporal variance. Thereafter, temporal encoding is an integral part of general perception. Moreover, time encoding on this level is an automatic process independent of attentional capacities (Golan & Zakay, 2015). Tobin and Grondin compared expert and intermediate runners to compare their running time with their predicted time. Results show that task experience affects temporal prediction and accuracy in actual running time estimation in the order of many minutes (Tobin & Grondin, 2015). Zhang and Chen showed that time perception is adaptively recalibrated and biased by quick statistical binding of temporal information and non-temporal stimuli properties, by using a visual Ternus display as probe (Zhang & Chen, 2016). Szelag et al. used temporal training protocols and explored the link between temporal information processing and language disorders (in aphasic patients and children with language impairment), and their therapy tools provide evidence for promising clinical applications (Szelag et al., 2015).

## Sensorimotor Synchronization, Embodiment and Coordination

Under this topic, we included four studies. Booth and Elliott investigated individuals' ability to synchronize movements to a temporal-spatial visual cue in the presence of same modality temporal-spatial distractors and found early but not late visual distractors affect movement synchronization to a temporal-spatial visual cue (Booth & Elliott, 2015). Hao et al. investigated the effect of voluntary movement on the simultaneous perception of auditory and tactile stimuli using a temporal order judgment task with voluntary movement, involuntary movement, and no movement, suggesting that the efference copy has a role in explaining the differential effects (Hao, Ogata, Ogawa, Kwon, & Miyake, 2015). In the framework of embodied time perception, Jia et al. showed that weight experience modulates visual duration estimation through the link



between the weight of the backpack and the to be estimated visual target (backpack picture) (Jia, Shi, & Feng, 2015).

Osaka et al. extended the investigation of time perception to two agents and examined how two brains make one synchronized behavior using cooperated singing/humming between two people and hyperscanning - a new brain scanning technique. They found a significant increase in neural synchronization of the left inferior frontal cortex (IFC) as a neural signature for cooperative singing or humming (Osaka et al., 2015).

### **Perspective of Psychological Moment and Temporal Organization**

The last part incorporates three papers which might provoke a re-thinking of concepts and methodology in sensory timing. Elliott and Giersch reconsidered the concept of “psychological moment” and suggested that within the 50-60 ms interval a more fine-scaled, serialized process structures and defines the passage of ongoing time. That is, a perceptual moment is experienced a co-temporality (two events are experienced as happening simultaneous) but on a level accessible through implicit behavioral measures nevertheless time is processed sequentially (Elliott & Giersch, 2015). Arstila discusses and then leans towards a brain time view, with respect to the debate of time-marker view vs. brain time view, a debate that is concerned with the question of how an observer extracts temporal information from a continuous stream of events (Arstila, 2015). Zhou et al. proposes that temporal aspects of objects can be treated as features of objects, and that psychological time or “apparent time”, similar to concepts underlying the analyses of reaction times, can serve as a tool to study the principles of neural codes related to object identity (Zhou, Zhang, & Mao, 2015).

Overall, the collections in “Sub-and supra-second timing: brain, learning and development” show some recent trends and debates in multisensory timing research as well as provide a venue to inspire future work in multisensory timing..

**Citation:** Chen, L., Bao, Y., Wittmann, M., eds. (2016). Sub- and Supra-Second Timing: Brain, Learning and Development. Lausanne: Frontiers Media. doi: 10.3389/978-2-88919-898-6

# Table of Contents

## **07 Editorial: Sub- and Supra-Second Timing: Brain, Learning and Development**

Lihan Chen, Yan Bao and Marc Wittmann

### **Sensory timing, interaction and reliability**

#### **10 Filling the blanks in temporal intervals: the type of filling influences perceived duration and discrimination performance**

Ninja K. Horr and Massimiliano Di Luca

#### **18 Duration estimates within a modality are integrated sub-optimally**

Ming Bo Cai and David M. Eagleman

#### **36 The influence of stimulus repetition on duration judgments with simple stimuli**

Teresa Birngruber, Hannes Schröter and Rolf Ulrich

#### **45 Visual-auditory differences in duration discrimination of intervals in the subsecond and second range**

Thomas H. Rammsayer, Natalie Borter and Stefan J. Troche

#### **52 Odors Bias Time Perception in Visual and Auditory Modalities**

Zhenzhu Yue, Tianyu Gao, Lihan Chen and Jiashuang Wu

### **Adaptive representation of time, learning and temporal prediction**

#### **61 The duality of temporal encoding – the intrinsic and extrinsic representation of time**

Ronen Golan and Dan Zakay

#### **72 Prior task experience affects temporal prediction and estimation**

Simon Tobin and Simon Grondin

#### **79 Crossmodal Statistical Binding of Temporal Information and Stimuli Properties Recalibrates Perception of Visual Apparent Motion**

Yi Zhang and Lihan Chen

#### **91 The Application of Timing in Therapy of Children and Adults with Language Disorders**

Elzbieta Szelag, Anna Dacewicz, Aneta Szymaszek, Tomasz Wolak, Andrzej Senderski, Izabela Domitrz and Anna Oron

#### **109 Corrigendum: The Application of Timing in Therapy of Children and Adults with Language Disorders**

Elzbieta Szelag, Anna Dacewicz, Aneta Szymaszek, Tomasz Wolak, Andrzej Senderski, Izabela Domitrz and Anna Oron



## **Sensorimotor synchronization, embodiment and coordination**

### **110 *Early, but not late visual distractors affect movement synchronization to a temporal-spatial visual cue***

Ashley J. Booth and Mark T. Elliott

### **118 *The simultaneous perception of auditory–tactile stimuli in voluntary movement***

Qiao Hao, Taiki Ogata, Ken-ichiro Ogawa, Jinhwan Kwon and Yoshihiro Miyake

### **128 *Wearing weighted backpack dilates subjective visual duration: the role of functional linkage between weight experience and visual timing***

Lina Jia, Zhuanghua Shi and Wenfeng Feng

### **135 *How Two Brains Make One Synchronized Mind in the Inferior Frontal Cortex: fNIRS-Based Hyperscanning During Cooperative Singing***

Naoyuki Osaka, Takehiro Minamoto, Ken Yaoi, Miyuki Azuma, Yohko Minamoto Shimada and Mariko Osaka

## **Perspective of psychological moment and temporal organization**

### **146 *What Happens in a Moment***

Mark A. Elliott and Anne Giersch

### **154 *Defense of the brain time view***

Valtteri Arstila

### **158 *Temporal perception in visual processing as a research tool***

Bin Zhou, Ting Zhang and Lihua Mao



# Editorial: Sub- and Supra-Second Timing: Brain, Learning and Development

Lihan Chen<sup>1,2\*</sup>, Yan Bao<sup>1,2,3</sup> and Marc Wittmann<sup>4</sup>

<sup>1</sup> Department of Psychology and Beijing Key Laboratory of Behavior and Mental Health, Peking University, Peking, China, <sup>2</sup> Key Laboratory of Machine Perception (Ministry of Education), Peking University, Beijing, China, <sup>3</sup> Human Science Center, Institute of Medical Psychology, Ludwig Maximilian University Munich, Munich, Germany, <sup>4</sup> Institute for Frontier Areas of Psychology and Mental Health, Freiburg, Germany

**Keywords:** subjective time, time perception, time coordination, movement timing, timing mechanisms

## The Editorial on the Research Topic

### Sub- and Supra-Second Timing: Brain, Learning and Development

Time perception in the range of milliseconds to a few seconds is essential for many important sensory and perceptual tasks including speech perception, motion perception, motor coordination, and cross-modal interaction. For the brain to be in synchrony with the environment, the physical differences in the speeds of light and sound, as well as stimuli from other modalities such as odors, must be processed and coordinated (Pöppel and Bao, 2014; Bao et al., 2015).

Time is a subjective feeling that is modulated by emotional states which trigger temporal distortions (temporal dilation vs. contraction; Wittmann, 2016), hence give rise to subjective time that may be different to *event time* as initially registered in the brain. Recent research suggests that time perception in a multisensory world is subject to prior task experience and shaped by (statistical) learning processes. Humans are active learners. That is, the engagement of the own body in a timing task within a perceptual-action loop will make a noticeable difference in timing performance, as compared to when humans only passively perceive the same perceptual scenario (Chen and Vroomen, 2013).

This Research Topic of “Sub- and supra-second timing: brain, learning and development” has integrated 16 submissions of novel research on sub- and supra-timing. We have categorized the papers in this topic into the following four themes, from which we can deduce trends of research about multisensory timing in the sub- and supra-second range.

## SENSORY TIMING, INTERACTION, AND RELIABILITY

A central debate in sensory timing is whether it is subserved by a centralized timing mechanism or distinctive/modular processing (Ivry and Schlerf, 2008). We included five papers underlying this theme. Di Luca investigated how judgments of perceived duration are influenced by the properties of the signals that define the intervals. They found that timing distortion is attributed to both intervals (isochronous vs. an-isochronous) and filling types (empty vs. filled) (Horr and Di Luca). Cai and Eagleman asked themselves how the brain forms a representation of duration when each of two simultaneously presented stimuli influence perceived duration in different ways. They attributed the perceived averaged duration of simultaneously occurring visual stimuli to the weightings of the elementary (individual) stimuli, although the weighting performance did not fully predict statistically optimal integration. Birngruber et al. examined the effects of stimulus repetition

## OPEN ACCESS

### Edited and reviewed by:

Rufin VanRullen,  
Centre de Recherche Cerveau et  
Cognition, France

### \*Correspondence:

Lihan Chen  
clh@pku.edu.cn

### Specialty section:

This article was submitted to  
Perception Science,  
a section of the journal  
Frontiers in Psychology

**Received:** 23 April 2016

**Accepted:** 06 May 2016

**Published:** 18 May 2016

### Citation:

Chen L, Bao Y and Wittmann M (2016)  
Editorial: Sub- and Supra-Second  
Timing: Brain, Learning and  
Development. *Front. Psychol.* 7:747.  
doi: 10.3389/fpsyg.2016.00747



vs. stimulus novelty on perceived duration. They substantiated the view that changes of simple, that is, semantically meaningless stimuli lead to shorter perceived duration of repeated as compared to novel stimuli. In the conceptual framework of the distinct timing hypothesis, Rammsayer et al. showed a gradual transition from a purely modality-specific, sensory-automatic to a more cognitive, a modal timing mechanism, by viewing the evidence that the prevalence of precision of auditory over visual timing disappeared when the temporal range is controlled. Yue et al. explored the effects of olfactory events upon reproduced time durations in auditory and visual modalities, and found that the biased timing in target stimuli (auditory and visual) could be accounted for by a framework of attentional deployment between the inducers (odors) and emotionally neutral stimuli (visual dots and sound beeps). However, the mechanisms of distinct timing vs. centralized timing are to be determined.

## ADAPTIVE REPRESENTATION OF TIME, LEARNING, AND TEMPORAL PREDICTION

Golan and Zakay probed the duality of temporal encoding—the intrinsic and extrinsic representation of time—using fMRI. They exposed participants to stimuli with different temporal variance and found neural activation (within category-selective brain regions) increase as a function of increase in temporal variance. Thereafter, temporal encoding is an integral part of general perception. Moreover, time encoding on this level is an automatic process independent of attentional capacities. Tobin and Grondin compared expert and intermediate runners to compare their running time with their predicted time. Results show that task experience affects temporal prediction and accuracy in actual running time estimation in the order of many minutes. Zhang and Chen showed that time perception is adaptively recalibrated and biased by quick statistical binding of temporal information and non-temporal stimuli properties, by using a visual Ternus display as probe. Szelag et al. used temporal training protocols and explored the link between temporal information processing and language disorders (in aphasic patients and children with language impairment), and their therapy tools provide evidence for promising clinical applications.

## SENSORIMOTOR SYNCHRONIZATION, EMBODIMENT, AND COORDINATION

Under this topic, we included four studies. Booth and Elliott investigated individuals' ability to synchronize movements to a temporal-spatial visual cue in the presence of same modality temporal-spatial distractors and found early but not late visual distractors affect movement synchronization

to a temporal-spatial visual cue. Hao et al. investigated the effect of voluntary movement on the simultaneous perception of auditory and tactile stimuli using a temporal order judgment task with voluntary movement, involuntary movement, and no movement, suggesting that the reference copy has a role in explaining the differential effects. In the framework of embodied time perception, Jia et al. showed that weight experience modulates visual duration estimation through the link between the weight of the backpack and the to be estimated visual target (backpack picture).

Osaka et al. extended the investigation of time perception to two agents and examined how two brains make one synchronized behavior using cooperated singing/humming between two people and hyperscanning—a new brain scanning technique. They found a significant increase in neural synchronization of the left inferior frontal cortex (IFC) as a neural signature for cooperative singing or humming.

## PERSPECTIVE OF PSYCHOLOGICAL MOMENT AND TEMPORAL ORGANIZATION

The last part incorporates three papers which might provoke a re-thinking of concepts and methodology in sensory timing. Elliott and Giersch reconsidered the concept of “psychological moment” and suggested that within the 50–60 ms interval a more fine-scaled, serialized process structures and defines the passage of ongoing time. That is, a perceptual moment is experienced as co-temporality (two events are experienced as happening simultaneous) but on a level accessible through implicit behavioral measures nevertheless time is processed sequentially. Arstila discusses and then leans toward a brain time view, with respect to the debate of *time-marker view* vs. *brain time view*, a debate that is concerned with the question of how an observer extracts temporal information from a continuous stream of events. Zhou et al. propose that temporal aspects of objects can be treated as features of objects, and that psychological time or “apparent time,” similar to concepts underlying the analyses of reaction times, can serve as a tool to study the principles of neural codes related to object identity.

Overall, the collections in “Sub- and supra-second timing: brain, learning and development” show some recent trends and debates in multisensory timing research as well as provide a venue to inspire future work in multisensory timing.

## AUTHOR CONTRIBUTIONS

LC drafted the editorial, YB and MW revised it.

## REFERENCES

- Bao, Y., Pöppel, E., Wang, L., Lin, X., Yang, T., Avram, M., et al. (2015). Synchronization as a biological, psychological and social mechanism to create common time: a theoretical frame and a single case study. *PsyCh J.* 4, 243–254. doi: 10.1002/pchj.119
- Chen, L., and Vroomen, J. (2013). Intersensory binding across space and time: a tutorial review. *Atten. Percept. Psychophys.* 75, 790–811. doi: 10.3758/s13414-013-0475-4
- Ivry, R. B., and Schlerf, J. E. (2008). Dedicated and intrinsic models of time perception. *Trends Cogn. Sci.* 12, 273–280. doi: 10.1016/j.tics.2008.04.002
- Pöppel, E., and Bao, Y. (2014). “Temporal windows as a bridge from objective to subjective time,” in *Subjective Time*, eds D. Lloyd and V. Arstila (Cambridge, MA: MIT Press), 241–261.
- Wittmann, M. (2016). *Felt Time: The Psychology of How We Perceive Time*. Cambridge, MA: MIT Press.

**Conflict of Interest Statement:** 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.

Copyright © 2016 Chen, Bao and Wittmann. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





# Filling the blanks in temporal intervals: the type of filling influences perceived duration and discrimination performance

Ninja K. Horr and Massimiliano Di Luca \*

Centre for Computational Neuroscience and Cognitive Robotics, School of Psychology, University of Birmingham, Birmingham, UK

**Edited by:**

Lihan Chen, Peking University, China

**Reviewed by:**

Zhuanghua Shi,  
Ludwig-Maximilians-Universität  
München, Germany  
Emi Hasuo, Kyushu University,  
Japan

**\*Correspondence:**

Massimiliano Di Luca, Centre for  
Computational Neuroscience and  
Cognitive Robotics, School of  
Psychology, University of  
Birmingham, Hills Building,  
Edgbaston, Birmingham B15 2TT,  
UK  
e-mail: m.diluca@bham.ac.uk

In this work we investigate how judgments of perceived duration are influenced by the properties of the signals that define the intervals. Participants compared two auditory intervals that could be any combination of the following four types: intervals filled with continuous tones (*filled* intervals), intervals filled with regularly-timed short tones (*isochronous* intervals), intervals filled with irregularly-timed short tones (*anisochronous* intervals), and intervals demarcated by two short tones (*empty* intervals). Results indicate that the type of intervals to be compared affects discrimination performance and induces distortions in perceived duration. In particular, we find that duration judgments are most precise when comparing two isochronous and two continuous intervals, while the comparison of two anisochronous intervals leads to the worst performance. Moreover, we determined that the magnitude of the distortions in perceived duration (an effect akin to the filled duration illusion) is higher for tone sequences (no matter whether isochronous or anisochronous) than for continuous tones. Further analysis of how duration distortions depend on the type of filling suggests that distortions are not only due to the perceived duration of the two individual intervals, but they may also be due to the comparison of two different filling types.

**Keywords:** temporal perception, perceived duration, short-interval duration, duration distortions, filled-duration illusion, interval filling

## INTRODUCTION

Many factors other than the physical duration of an interval influence perceived duration (see Allan, 1979 for a classic and Grondin, 2010 for a recent overview). For example, perceived duration is influenced by the filling of the interval to be judged as highlighted by the well-known filled duration illusion, whereby filled intervals are perceived as longer than their empty counterparts. This effect has been observed in a wide range of experimental conditions, with the definition of “filling” varying across studies. Several studies used continuous signals as filled intervals (e.g., Goldfarb and Goldstone, 1963; Steiner, 1968; Craig, 1973; Wearden et al., 2007; Hasuo et al., 2014) and compared those to empty intervals, which are typically consisting solely of a short beginning and end marker or a gap in a continuous signal (see Wearden et al., 2007 for a comparison of those two variations). Another type of filled interval leading to the filled duration illusion is a sequence of short filler signals that is compared to an empty interval lacking such fillers (e.g., Buffardi, 1971; Thomas and Brown, 1974; Adams, 1977). The magnitude of the overestimation for the latter type of filled intervals has been shown to increase with the number of fillers (Buffardi, 1971; Schiffman and Bobko, 1977). This overestimation has been termed “Illusion of a Divided Time Interval” by ten Hoopen et al. (2008).

Duration judgments with filled intervals are mostly investigated with regularly-timed tones—that is, isochronous rhythms.

However, it has recently been reported that the temporal structure of fillers influences perceived duration. For example, Matthews (2013) showed that isochronous intervals are perceived to last longer than accelerating or decelerating ones. Horr and Di Luca (2014) found that isochronous intervals are perceived to last longer than anisochronous ones and that this effect increases not only with the amount of anisochrony but also, like the filled duration illusion, with the number of fillers (this is in accordance with tendencies found in earlier studies, see Grimm, 1934; Thomas and Brown, 1974).

Overall, this line of research indicates that the type and structure of interval filling influences perceived duration. To gain further insight into the mechanisms underlying short interval duration perception also discrimination performance has to be investigated experimentally. Rammsayer and Lima (1991) reported that filled intervals made up of a continuous signal are discriminated better than empty intervals. It remains to be determined, whether this superior discrimination of filled as compared to empty intervals is only true for one type of filled intervals, namely intervals filled with a continuous signal (e.g., a continuous sound) or can as well be generalized over intervals filled with sequences of short filler signals (e.g., short tones). I further remains to be investigated how discrimination performance differs between such continuous and short filler intervals of different temporal structure.

In the present article we investigate how the type of interval filling affects perceived duration and discrimination performance using four types of auditory intervals: continuous, isochronous, anisochronous, and empty intervals. In Experiment 1 we investigate duration discrimination performance by having participants compare two intervals of the same type. In Experiment 2 we aim at quantifying the perceptual distortions for each interval type. To our knowledge, this is the first attempt to quantify how the type of filling influences the magnitude of the “filled duration illusion.” Such discrimination is important to understand the mechanisms involved in short-interval duration perception as it constrains the type of cognitive mechanisms employed in prospective time judgments.

## GENERAL METHODS

### PARTICIPANTS

A total of 35 healthy volunteers with normal auditory sensitivity participated in the experiments for course credits or a payment of 7 GBP/h. All participants were naive to the purpose of the study, reported normal auditory sensitivity and took part in only one of the experiments. The experimental data collection and storage followed the ethical guidelines of the Declaration of Helsinki and was approved by the Science, Technology, Engineering, and Mathematics Ethical Review Committee of the University of Birmingham.

### EXPERIMENTAL DESIGN

Participants performed a two-interval forced-choice task, deciding via button pressing which of two intervals had been the one of longer duration. A trial consisted of a 1000 ms standard interval and a comparison interval of 500, 700, 850, 1000, 1150, 1300, or 1500 ms duration spaced by a random interval between 2000 and 2300 ms. The order of standard and comparison intervals was random and counterbalanced across trials. Experimental stimuli constituting an interval were 1000 Hz 70 dB tones with 2.5 ms ramped onset and offset. Each interval consisted either of (a) a beginning and end tone lasting for 10 ms each (empty interval), (b) five 10 ms regularly-timed filler tones (isochronous interval), (c) five 10 ms irregularly-timed filler tones (anisochronous interval) or of (d) a tone lasting for the entire interval duration (continuous interval). For the anisochronous intervals, temporal irregularity was created by randomly moving the onset of individual filler tones inside a range of plus or minus half of the interstimulus interval (i.e., 250 ms in the standard interval). Stimuli were presented via headphones. Participants' individual response proportions were assessed in relation to the physical duration difference between interval types. The point of subjective equality (PSE) and the just noticeable difference (JND) were estimated using the Spearman-Kärber-Method as the first and second moment of the data obtained from each participant (Ulrich and Miller, 2004).

## EXPERIMENT 1: DURATION DISCRIMINATION PERFORMANCE

To investigate differences in duration discrimination performance across interval types, we asked participants to compare

two intervals of the same type (continuous, isochronous, anisochronous and empty).

## MATERIALS AND METHODS

Seventeen healthy volunteers (15 female,  $21.7 \pm 2.8$  years) participated in Experiment 1. In each experimental trial, participants reported which of two intervals was longer. According to the different interval types, four conditions were defined: continuous, isochronous, anisochronous, and empty. Each of the four conditions was presented in a block. The sequences of blocks (conditions) were randomized for each participant. Every block contained eight repetitions of all seven possible durations of the comparison interval (Mayer et al., 2014). In every block the eight repetitions of each comparison duration were counterbalanced and pseudo-randomized according to which interval (standard or comparison) was presented first. In total participants made 224 duration comparisons in 4 blocks of 56 trials each. The entire experiment lasted about 40 min.

## RESULTS

In **Figure 1A** response proportions and **Figure 1B** PSE and JND values are displayed. Each participant's average JND is lower than 600 ms, which means that all of them were reasonably capable of performing the task. As participants were comparing two identical intervals, there should be no difference between PSE values across conditions [ $F_{(3, 67)} = 1.6$ , n.s.]. More interestingly, there is a significant difference of JND values between conditions [ $F_{(3, 67)} = 15.4$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.49$ ].

*Post-hoc* tests indicate that the following differences are statistically significant: Duration discrimination is better for continuous than empty [paired sample *t*-test on JND,  $t_{(16)} = 3.9$ ,  $p = 0.0013$ ] and anisochronous intervals [ $t_{(16)} = 7.6$ ,  $p < 0.001$ ]. Discrimination is better for isochronous than empty [ $t_{(16)} = -2.2$ ,  $p = 0.043$ ] and anisochronous intervals [ $t_{(16)} = 4.5$ ,  $p < 0.001$ ]. Furthermore, discrimination is better for empty than anisochronous intervals [ $t_{(16)} = 2.4$ ,  $p = 0.03$ ]. There is no significant difference between continuous and isochronous intervals [ $t_{(16)} = 1.7$ ,  $p = 0.12$ ]. In short, continuous and isochronous intervals are discriminated best, followed by empty intervals, while discrimination performance is worst for anisochronous intervals.

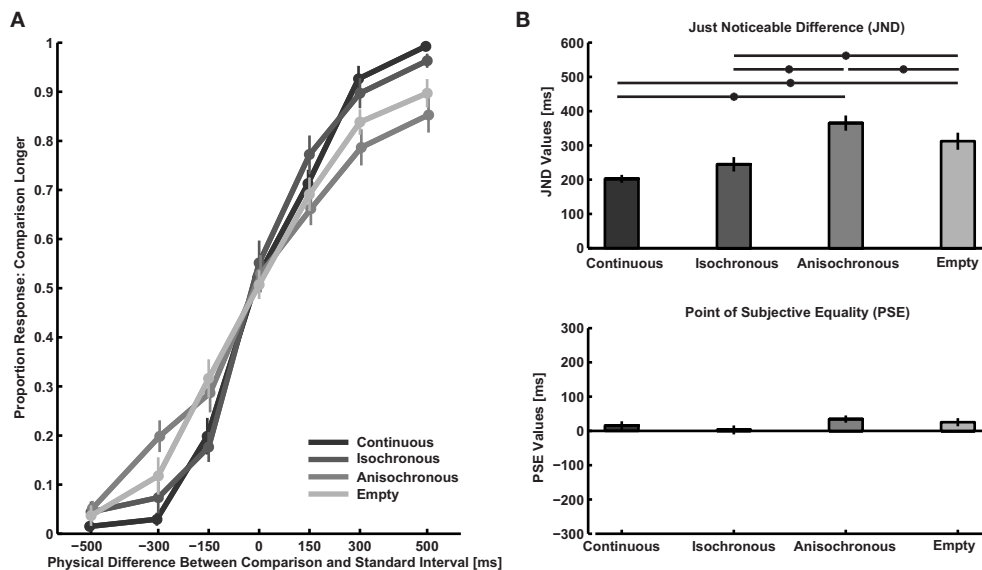
## EXPERIMENT 2: DISTORTIONS OF PERCEIVED DURATION

To investigate whether distortions of perceived duration depend on the type of interval filling, we asked participants to compare perceived duration between all types of filled intervals and the empty intervals. Furthermore, we asked participants to compare the duration of different types of filled intervals.

## MATERIAL AND METHODS

Eighteen healthy volunteers (12 female,  $22.1 \pm 3.3$  years) participated in Experiment 2. In each trial, participants made their duration judgment for two intervals of different types. Six conditions were defined according to all possible combinations of the four interval types: (1) continuous/empty, (2) isochronous/empty, (3) anisochronous/empty, (4) continuous/isochronous, (5) continuous/anisochronous, and (6) isochronous/anisochronous. Each condition was presented in a separate block of trials. As in





**FIGURE 1 | Results of Experiment 1. (A)** Participants' responses recorded to indicate the proportion of responses where the comparison interval was judged longer than the standard interval as a function of physical duration difference. **(B)** Point of subjective equality (PSE) and just

noticeable difference (JND) calculated from response proportions using the Spearman-Kärber method. Asterisks indicate differences in performance between intervals of different types as identified by the horizontal lines. Error bars are S.E.M.

Experiment 1 sequences of blocks (conditions) and trials were fully randomized. The order of standard (1000 ms) and comparison (500–1500 ms) intervals was counterbalanced and the standard could be either of the two types of intervals presented in the block. Data from the combination of order and standard type is presented combined. Participants performed a total of 336 duration discrimination judgments resulting from 6 blocks of 56 trials each. The entire experiment lasted about 60 min.

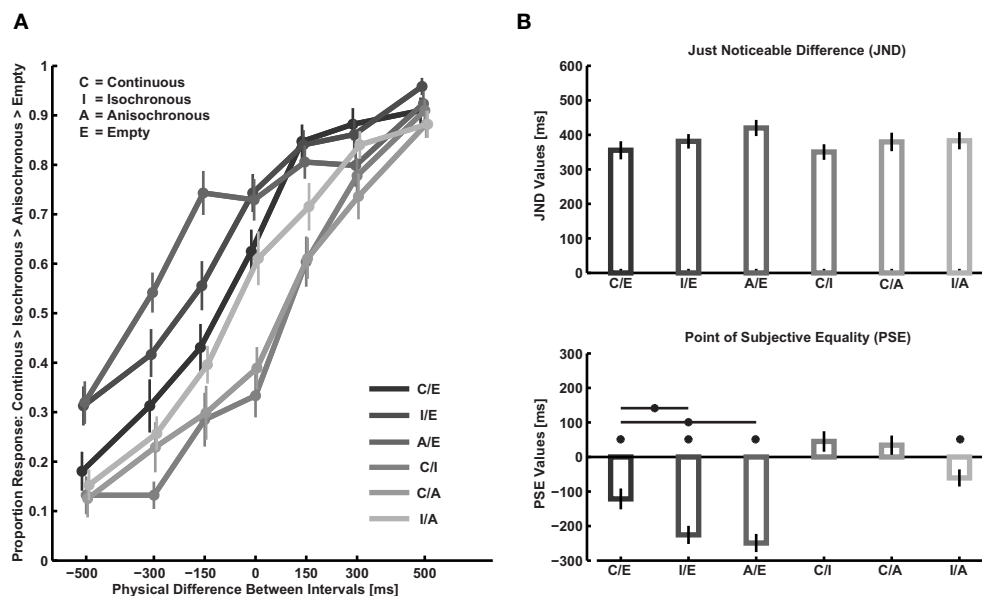
## RESULTS

**Figure 2A** shows response proportions and **Figure 2B** shows average PSE and JND values obtained across participants. Again as in Experiment 1, average JND values for each participant are lower than 600 ms indicating a reasonable performance. The PSE values depend on the type of filling [One-Way r.m. ANOVA:  $F_{(5,107)} = 23.4$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.58$ ]. In every conditions containing empty intervals PSEs are significantly lower than zero [single sample  $t$ -test on PSE against 0, continuous/empty:  $t_{(17)} = -4.0$ ,  $p < 0.001$ ; isochronous/empty:  $t_{(17)} = -8.6$ ,  $p < 0.001$ ; anisochronous/empty:  $t_{(17)} = -9.4$ ,  $p < 0.001$ ]. This indicates the presence of the filled duration illusion, that is, the duration of empty intervals being underestimated as compared to filled intervals. Isochronous intervals are perceived as longer than anisochronous ones [ $t_{(17)} = -2.5$ ,  $p = 0.025$ ], whereas PSE does not differ from 0 when comparing continuous and isochronous [ $t_{(17)} = 1.5$ ,  $p = 0.15$ ] as well as continuous and anisochronous intervals [ $t_{(17)} = 1.2$ ,  $p = 0.24$ ]. The magnitude of bias (PSE value) is lower for continuous intervals than for isochronous intervals [paired sample  $t$ -test on PSE isochronous/empty vs. PSE continuous/empty:  $t_{(17)} = 3.0$ ,  $p = 0.008$ ] as well as for anisochronous intervals [PSE anisochronous/empty vs. PSE continuous/empty:  $t_{(17)} =$

3.5,  $p = 0.003$ ]. There is no significant difference between isochronous and anisochronous [PSE isochronous/empty vs. PSE anisochronous/empty  $t_{(17)} = 0.8$ ,  $p = 0.43$ ]. No significant difference is observed in JND values across conditions [One-Way ANOVA on JND,  $F_{(5,107)} = 2.0$ ,  $p = 0.09$ ,  $\eta_p^2 = 0.10$ ], with a tendency toward better performance in conditions where one of the compared stimuli is a continuous interval. A comparison of JND values between Experiment 1 and 2 indicates higher performance when comparing intervals of the same type rather than of different types [two sample  $t$ -test on average JND for each participant:  $t_{(33)} = 4.3$ ,  $p < 0.001$ ,  $0.38 \pm 0.02$  ms vs.  $0.28 \pm 0.01$  ms].

## GENERAL DISCUSSION

The present article investigates discrimination performance and perceived duration of four types of auditory intervals: continuous tones, isochronous sequences of tones, anisochronous sequences of tones, and empty intervals. Such interval types have been commonly used in experiments investigating the filled duration illusion and related distortions of perceived duration (e.g., Thomas and Brown, 1974; Rammsayer and Lima, 1991; Wearden et al., 2007), but until now they have never been systematically compared. We find that discrimination performance changes depending on the interval types to be compared. When comparing the same types of intervals, continuous and isochronous intervals are discriminated better than empty intervals. Discrimination performance for anisochronous intervals is worse than for all other interval types. The filled duration illusion is found to be stronger for tone sequences, both isochronous and anisochronous, than for continuous intervals. The result of the comparison of different types of filled intervals, however, indicates that there are no differences in duration judgments between continuous tones and



**FIGURE 2 | Results of Experiment 2. (A)** Proportions of judging continuous > isochronous > anisochronous > empty as a function of duration difference between standard and comparison. **(B)** Point of subjective equality (PSE) and just noticeable difference (JND) calculated

from response proportions using the Spearman-Kärber method. Asterisks indicate a significant difference of the PSE from zero and between the three conditions comprising one empty interval. Error bars are S.E.M.

tone sequences, and that isochronous sequences are perceived as longer than anisochronous ones.

### DISCRIMINATION PERFORMANCE

Differences in duration discrimination performance between interval types demonstrate that participants make use of the structure of interval filling to arrive at their duration estimates. That is, for the different interval types they use either different sources of information or there is a common mechanism that changes in precision depending on the interval types.

Our data indicates that when comparing intervals of the same type, continuous and isochronous intervals are better discriminated than empty ones. This is in line with the idea that higher sound energy in the interval improves discrimination performance (Carbotte and Kristofferson, 1973). However, empirical evidence that do not support this possibility (Creelmann, 1962; Abel, 1972). Rammsayer and Lima (1991) suggest that filled intervals are discriminated better than empty intervals because they elicit a higher neural firing rate, which is translated to a superior temporal resolution. This possibility would predict a better discrimination performance for sound sequences than for continuous intervals because a continuous sound would be subject to habituation (e.g., Polich, 1989). In addition, Horr and Di Luca (2014) hypothesized that due to neural entrainment (e.g., Lakatos et al., 2008; Cravo et al., 2013), stimuli in isochronous sequences should arrive at the point of highest neural responsiveness leading to further increase in neural response in isochronous intervals when compared to continuous intervals. However, our results (Figure 1B) do not show a significant difference between continuous intervals and isochronous sequences. Also the finding of anisochronous sequences being discriminated worse than

continuous tones and empty intervals is not in accordance with a neural firing rate explanation. The higher temporal resolution caused by increased neural responses can therefore only account for the decrease in performance found with empty as compared to continuous and isochronous intervals, as the lack of difference between continuous and isochronous intervals and even more so the remarkably worse performance for anisochronous as compared to all other intervals remains unexplained.

Another possibility to explain the observed pattern of discrimination performance is to appeal to the number of cues available for a single duration judgment. It has been shown that filled intervals defined by auditory and visual stimuli provide redundant cues to duration that allow a statistically optimal increase in performance (Hartcher-O'Brien et al., 2014). Here we posit that in some conditions there are redundant cues related to duration also for unisensory stimuli and this could lead to better discrimination performance compared to the conditions where only one cue is available. In particular, Hartcher-O'Brien et al. (2014) identify the filling of the interval as an important factor that can modulate the modality of integration, as empty intervals consist of two markers that only allow the identification of two time points and of the subtended empty duration between them. In contrast, continuous tones allow duration estimates by using the overall sensed energy in addition to (and independently from) the information carried by the temporal difference between beginning and ending time points. For isochronous intervals, the regular temporal structure allows to estimate duration based solely on the interval between successive tones (if the number of tones is known). Although the same cue is present with anisochronous intervals, the random timing of tones should be actually deceptive and lead to a reduced precision in duration judgments. If we interpret our

data along these lines, the pattern of results suggests that the base duration judgment performance is achieved with empty intervals. In filled intervals the brain can use additional duration cues if both intervals carry such cues, that is, with trials with two intervals of the same type as in our Experiment 1. Such cues can either increase (as in the case of isochronous intervals), but also decrease discrimination performance (as with anisochronous intervals). If two intervals of different types are compared, additional cues cannot be used, leading to a lower discrimination performance in all conditions in Experiment 2. Such cues are present while comparing anisochronous intervals but they decrease rather than increase discrimination performance. On the other hand, such cues cannot be compared directly with stimuli of different types, leading to lower discrimination performance in all conditions of Experiment 2.

### DISTORTIONS OF PERCEIVED DURATION

The goal of Experiment 2 was to characterize duration distortions. PSE data shows that the effect of the filled duration illusion (e.g., Steiner, 1968; Buffardi, 1971; Thomas and Brown, 1974; Wearden et al., 2007; Hasuo et al., 2014) is present for every type of filled interval we tested. The data however indicates that the magnitude of the filled duration illusion is higher with isochronous and anisochronous than with continuous intervals. That is, PSE values are significantly lower for the comparison between isochronous/empty and anisochronous/empty than for continuous/empty intervals. We hypothesize that different additional duration cues present in filled intervals could be responsible for this. For example, for some comparison types participants could use neural response magnitudes, as there seems to be a positive relation between those and perceived duration (see Eagleman and Pariyadath, 2009). The difference in the results with continuous intervals and tone sequences could then be due to the comparatively lower neural response with continuous intervals due to neural adaptation (e.g., Polich, 1989). The higher peak of neural response with isochronous as compared to continuous intervals could further be due to neural entrainment, at the expected time points (Lakatos et al., 2008). Appealing to overall energy in neural responses is intriguing because it can account for the filled duration illusion, for the higher effect of tone sequences as compared to continuous tone and for the here replicated difference between isochronous and anisochronous intervals (Horr and Di Luca, 2014). An alternative explanation for the differentiation between isochronous and anisochronous intervals taken alone could be a logarithmic relationship between physical and perceived duration of intervals between tones (see Thomas and Brown, 1974; Matthews, 2013; Horr and Di Luca, 2014).

The attempt to account for the overall pattern of results in Experiment 2 by appealing to one of the discussed single mechanism is limited by two apparent internal inconsistencies of the data. (1) Even though the direct comparison of isochronous with anisochronous intervals leads to a noticeable difference in perceived duration, the magnitude of the filled duration illusion measured by comparing a filled to an empty interval is *not* different for isochronous as compared to anisochronous intervals. (2) Even though the direct comparison of tone sequences (both isochronous and anisochronous) with continuous intervals

does not lead to a significant difference, the filled duration illusion is weaker for continuous sounds than for isochronous and anisochronous intervals (again measured by comparing a filled to an empty interval).

To investigate the magnitude of inconsistencies in our data, we used the PSE values from the different comparison conditions to calculate relative duration distortions for each interval type as described in Mayer et al. (2014). Here we can express PSE values as the difference in the two physical durations  $PSE_{12} = D_1 - D_2$  that leads to identical perceived durations  $D'_1 = D'_2$ . As perceived duration can be expressed as  $D' = D + \widetilde{d}$ , where  $\widetilde{d}$  represents the distortions in perceived duration  $D'$  from the objective duration  $D$ , we can formulate PSE as a function of perceived durations and distortions:

$$PSE_{12} = D_1 - D_2 = D'_1 - \widetilde{d}_1 - D'_2 + \widetilde{d}_2.$$

But because perceived durations  $D'_1$  and  $D'_2$  are identical at PSE, we can simplify the formula as the difference in duration distortion:

$$PSE_{12} = D_1 - D_2 = \widetilde{d}_2 - \widetilde{d}_1.$$

In fact, PSE can be expressed not only relatively to the objective duration  $D$ , but also as the difference in duration distortion  $d$  from any value  $a$  as such:

$$PSE_{12} = \left( a + \widetilde{d}_2 \right) - \left( a + \widetilde{d}_1 \right) = \widetilde{d}_2 - \widetilde{d}_1.$$

In the following,  $d_1$  and  $d_2$  will represent the relative distortion in perceived duration with respect to  $a$ , the average duration distortion in the experiment. If we want to express the six PSEs obtained in the conditions of Experiment 2, we can use the following system of equations:

$$\begin{bmatrix} pse_{continuous/empty} \\ pse_{isochronous/empty} \\ pse_{anisochronous/empty} \\ pse_{continuous/isochronous} \\ pse_{continuous/anisochronous} \\ pse_{isochronous/anisochronous} \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 & 1 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 1 \\ -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 \end{bmatrix} \begin{bmatrix} d_{continuous} \\ d_{isochronous} \\ d_{anisochronous} \\ d_{empty} \end{bmatrix},$$

that is:

$$p = M d.$$

If  $d$  were the absolute value of distortion, such system would have infinite solutions. But here we express  $d$  relatively to the average duration distortion in the experiment  $a$ , so that a single solution to this linear system can be approximated using the Moore-Penrose pseudoinverse  $M^+$ :

$$d_{estimated} = M^+ p.$$

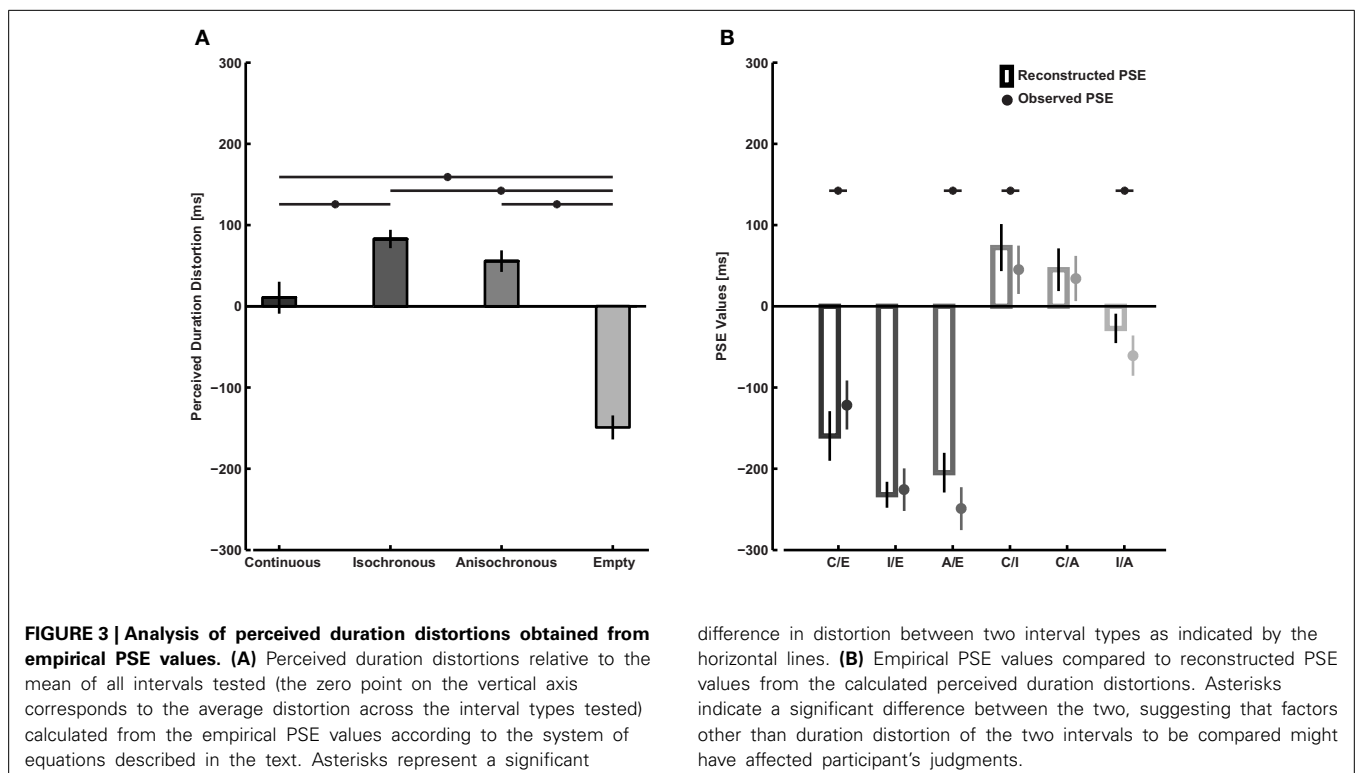
We apply this formula to the data obtained from each participant so to calculate the mean distortion in perceived duration for the four types of intervals tested (**Figure 3A**). Here,  $d = 0$  refers to a duration distortion equal to the average duration distortion  $a$  over all interval types tested in Experiment 1 (see Mayer et al., 2014). Empty intervals are perceived as shorter than continuous intervals [paired sample  $t$ -test on  $d$  values,  $t_{(17)} = 5.2$ ,  $p < 0.001$ ], isochronous intervals [ $t_{(17)} = 14.5$ ,  $p < 0.001$ ], and anisochronous intervals [ $t_{(17)} = 8.4$ ,  $p < 0.001$ ]. Moreover, continuous intervals are perceived as shorter than isochronous ones [ $t_{(17)} = -2.5$ ,  $p = 0.02$ ]. There is no difference between continuous vs. anisochronous [ $t_{(17)} = -1.7$ ,  $p = 0.10$ ] nor isochronous vs. anisochronous [ $t_{(17)} = 1.5$ ,  $p = 0.15$ ] intervals. Reconstructing the PSE from calculated distortions is possible using:

$$P_{reconstructed} = M d_{estimated}.$$

Such formula makes it possible to determine whether PSE values in the comparison task were solely dependent on the sum of single interval distortions. The comparison between observed and reconstructed PSE values is displayed in **Figure 3B**. Observed and reconstructed data differ significantly as indicated by the interaction term of a Two-Way r.m. ANOVA on PSE values with factors condition and empirical/reconstructed [ $F_{(5,85)} = 5.3$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.24$ ]. The values for the continuous/empty [paired sample  $t$ -test on PSE,  $t_{(17)} = 2.8$ ,  $p = 0.013$ ], anisochronous/empty [ $t_{(17)} = -3.4$ ,  $p = 0.003$ ], continuous/isochronous [ $t_{(17)} = -2.7$ ,  $p = 0.016$ ], and isochronous/anisochronous conditions [ $t_{(17)} = -2.7$ ,  $p = 0.015$ ]

differ significantly between empirical and reconstructed. Only the difference in the continuous/anisochronous [ $t_{(17)} = -0.9$ ,  $p = 0.36$ ] and isochronous/empty conditions [ $t_{(17)} = 0.47$ ,  $p = 0.64$ ] were not significant.

These inconsistencies indicate that distortions in two-interval forced-choice duration judgments do not solely depend on the perceived duration of the two intervals compared, which challenges the assumption of simple difference models (see e.g., Green and Swets, 1973; Thurston, 1994; Macmillan and Creelman, 2005). Context effects regarding the sequence in which stimuli are presented (e.g., Hellström, 1985, 2003; Dyjas and Ulrich, 2014) and the distribution of durations (e.g., Wearden and Ferrara, 1995; Brown et al., 2005; Wearden and Lejeune, 2008; Jazayeri and Shadlen, 2010) have frequently been reported in the literature. To test whether our results could be accounted for by hysteresis in duration judgments, i.e., if there is a distortion of perceived duration depending on the type of filling of the previous interval, we performed a  $2 \times 6$  Two-Way r.m. ANOVA on PSE values with factors presentation order (which of the two intervals was presented first) and comparison type (the six comparison conditions, cf. **Figure 2**). In accordance with the literature (e.g., Hellström, 2003; Dyjas and Ulrich, 2014) we find a significant bias to judge the second interval as longer than the first one [ $F_{(1, 17)} = 12.7$ ,  $p = 0.002$ ,  $\eta_p^2 = 0.57$ ] and as expected the factor comparison type is significant [ $F_{(5, 85)} = 23.45$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.43$ ]. Most importantly there is no significant interaction between the two factors order and comparison type [ $F_{(5, 85)} = 1.94$ , n.s.] suggesting that the inconsistencies in PSE we found cannot be accounted for by appealing to the presentation order of the intervals alone.





It thus remains unclear what are the factors inducing inconsistencies in the data across conditions, but one may speculate that different mechanisms could be used to compare durations when intervals to be compared are of the same type and of different type. We have discussed previously that duration judgments performed with the same type of intervals as in Experiment 1 could be aided by additional cues that are correlated to temporal duration (i.e., total energy and timing between successive tones). With the exception of isochronous and anisochronous intervals, the trials in Experiment 2 do not allow a direct comparison of additional cues to duration. Participants may have tried to map different cues to improve the comparison (i.e., mapping total energy in one interval to subinterval duration) thus creating response biases leading to one type of interval to be reported longer more often than the other (irrespective of the physical duration). Such biases are dependent on the pair of stimuli involved in the comparison and could thus explain the inconsistencies we observed in our data.

## CONCLUSIONS

Our results highlight the influence of interval type on discrimination performance and perceived duration. The observed effects have several implications regarding the computational and neural mechanisms underlying duration judgments. Differences in discrimination performance can be explained by considering the presence of multiple cues for duration discrimination when comparing intervals of the same type. Also distortions in perceived duration can be accounted for by appealing to such additional cues, particularly neural response magnitude, which is higher for continuous and anisochronous stimuli compared to empty, but is even higher with isochronous stimuli due to neural entrainment. Interestingly, inconsistencies in the pattern of results indicate that duration judgments in a forced-choice comparison task are affected by factors other than distortions in perceived duration of the individual intervals. Such factors need to be taken into account to understand internal inconsistencies in duration comparisons between different interval types.

## ACKNOWLEDGMENTS

This research was funded by the Marie Curie Grant CIG 304235 “TICS.”

## REFERENCES

- Abel, S. M. (1972). Duration discrimination of noise and tone bursts. *J. Acoust. Soc. Am.* 51, 1219–1223. doi: 10.1121/1.1912963
- Adams, R. D. (1977). Intervening stimulus effects on category judgments of duration. *Percept. Psychophys.* 21, 527–534. doi: 10.3758/BF03198733
- Allan, L. G. (1979). The perception of time. *Percept. Psychophys.* 26, 340–354. doi: 10.3758/BF03204158
- Brown, G. D., McCormack, T., Smith, M., and Steward, N. (2005). Identification and bisection of temporal durations and tone frequencies: common models for temporal and nontemporal stimuli. *J. Exp. Psychol. Hum. Percept. Perform.* 31, 354–362. doi: 10.1037/0096-1523.31.5.919
- Buffardi, L. (1971). Factors affecting the filled-duration illusion in the auditory, tactual, and visual modalities. *Percept. Psychophys.* 10, 292–294. doi: 10.3758/BF03212828
- Carbotte, R. M., and Kristofferson, A. B. (1973). On energy-dependent cues in duration discrimination. *Percept. Psychophys.* 14, 501–505. doi: 10.3758/BF03211190
- Craig, J. C. (1973). A constant error in the perception of brief temporal intervals. *Percept. Psychophys.* 13, 99–104. doi: 10.3758/BF03207241
- Cravo, A. M., Rohenkohl, G., Wyart, V., and Nobre, A. C. (2013). Temporal expectation enhances contrast sensitivity by phase entrainment of low-frequency oscillations in visual cortex. *J. Neurosci.* 33, 4002–4010. doi: 10.1523/JNEUROSCI.4675-12.2013
- Creelmann, C. D. (1962). Human discrimination of auditory duration. *J. Acoust. Soc. Am.* 34, 582–593. doi: 10.1121/1.1918172
- Dyjas, O., and Ulrich, R. (2014). Effect of stimulus order on discrimination processes in comparative and equality judgements: data and models. *Q. J. Exp. Psychol.* 67, 1121–1150. doi: 10.1080/17470218.2013.847968
- Eagleman, D. M., and Pariyadath, V. (2009). Is subjective duration a signature of coding efficiency? *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 1841–1851. doi: 10.1098/rstb.2009.0026
- Goldfarb, J. L., and Goldstone, S. (1963). Time judgment: a comparison of filled and unfilled durations. *Percept. Mot. Skills* 16, 1963. doi: 10.2466/pms.1963.16.2.376
- Green, D. M., and Swets, J. A. (1973). *Signal Detection Theory and Psychophysics*. Huntington, NY: Krieger Publishing.
- Grimm, K. (1934). Der Einfluß der Zeitform auf die Wahrnehmung der Zeitdauer. *Z. Psychol.* 132, 104–132.
- Grondin, S. (2010). Timing and time perception: a review of recent behavioral and neuroscience findings. *Atten. Percept. Psychophys.* 72, 561–582. doi: 10.3758/APP.72.3.561
- Hartcher-O'Brien, J., Di Luca, M., and Ernst, M. O. (2014). The duration of uncertain times: audiovisual information about intervals is integrated in a statistically optimal fashion. *PLoS ONE* 9:e89339. doi: 10.1371/journal.pone.0089339
- Hasuo, E., Nakajima, Y., Tomimatsu, E., Grondin, S., and Ueda, K. (2014). The occurrence of the filled duration illusion: a comparison of the method of adjustment with the method of magnitude estimation. *Acta Psychol. (Amst.)* 147, 111–121. doi: 10.1016/j.actpsy.2013.10.003
- Hellström, A. (1985). The time-order error and its relatives-mirrors of cognitive processes in comparing. *Psychol. Bull.* 97, 35–61. doi: 10.1037/0033-2909.97.1.35
- Hellström, A. (2003). Comparison is not just subtraction: effects of time- and space-order on subjective stimulus difference. *Percept. Psychophys.* 65, 1161–1177. doi: 10.3758/BF03194842
- Horr, N. K., and Di Luca, M. (2014). Taking a long look at isochrony: perceived duration increases with temporal, but not stimulus regularity. *Atten. Percept. Psychophys.* doi: 10.3758/s13414-014-0787-z. [Epub ahead of print].
- Jazayeri, M., and Shadlen, M. N. (2010). Temporal context calibrates interval timing. *Nat. Neurosci.* 13, 1020–1026. doi: 10.1038/nn.2590
- Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., and Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science* 320, 110–113. doi: 10.1126/science.1154735
- Macmillan, N. A., and Creelman, C. D. (2005). *Detection Theory: A User's Guide*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Matthews, W. J. (2013). How does sequence structure affect the judgment of time? Exploring a weighted sum of segments model. *Cogn. psychol.* 66, 259–282. doi: 10.1016/j.cogpsych.2013.01.001
- Mayer, K. M., Di Luca, M., and Ernst, M. O. (2014). Duration perception in crossmodally-defined intervals. *Acta Psychol.* 147, 2–9. doi: 10.1016/j.actpsy.2013.07.009
- Polich, J. (1989). Habituation of P300 from auditory stimuli. *Psychobiology* 17, 19–28.
- Rammesayer, T. H., and Lima, S. D. (1991). Duration discrimination of filled and empty auditory intervals: cognitive and perceptual factors. *Percept. Psychophys.* 50, 565–574. doi: 10.3758/BF03207541
- Schiffman, H. R., and Bobko, D. J. (1977). The role of number and familiarity of stimuli in the perception of brief temporal intervals. *Am. J. Psychol.* 90, 85–93. doi: 10.2307/1421643
- Steiner, S. (1968). Apparent duration of auditory stimuli. *J. Aud. Res.* 8, 195–205.
- ten Hoopen, G., Miyauchi, R., and Nakajima, Y. (2008). “Time-based illusions in the auditory mode,” in *Psychology of Time*, ed S. Grondin (Bingley: Emerald Group), 139–188.
- Thomas, E. C., and Brown, I. (1974). Time perception and the filled-duration illusion. *Percept. Psychophys.* 16, 449–458. doi: 10.3758/BF03198571
- Thurston, L. L. (1994). A law of comparative judgment. *Psychol. Rev.* 101, 266–270. doi: 10.1037/0033-295X.101.2.266
- Ulrich, R., and Miller, J. (2004). Threshold estimation in two-alternative forced-choice (2AFC) tasks: the Spearman-Kärber method. *Percept. Psychophys.* 66, 517–533. doi: 10.3758/BF03194898

- Wearden, J. H., and Ferrara, A. (1995). Stimulus spacing effects in temporal bisection by humans. *Q. J. Exp. Psychol.* 48, 289–310.
- Wearden, J. H., and Lejeune, H. (2008). Sclar properties in human timing: conformity and violations. *Q. J. Exp. Psychol.* 61, 1875–1908. doi: 10.1080/17470210701282576
- Wearden, J. H., Norton, R., Martin, S., and Oliver, M.-B. (2007). Internal clock processes and the filled-duration illusion. *J. Exp. Psychol. Hum. Percept. Perform.* 33, 716–729. doi: 10.1037/0096-1523.33.3.716

**Conflict of Interest Statement:** 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.

*Received: 11 December 2014; accepted: 22 January 2015; published online: 11 February 2015.*

*Citation: Horr NK and Di Luca M (2015) Filling the blanks in temporal intervals: the type of filling influences perceived duration and discrimination performance. Front. Psychol. 6:114. doi: 10.3389/fpsyg.2015.00114*

*This article was submitted to Perception Science, a section of the journal Frontiers in Psychology.*

*Copyright © 2015 Horr and Di Luca. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

# Duration estimates within a modality are integrated sub-optimally

Ming Bo Cai and David M. Eagleman\*

Laboratory for Perception and Action, Department of Neuroscience, Baylor College of Medicine, Houston, TX, USA

## OPEN ACCESS

### Edited by:

Lihan Chen,  
Peking University, China

### Reviewed by:

Hansem Sohn,  
Massachusetts Institute of  
Technology, USA  
Alan Johnston,  
University College London, UK

### \*Correspondence:

David M. Eagleman,  
Laboratory for Perception and Action,  
Department of Neuroscience, Baylor  
College of Medicine, One Baylor Plaza  
Room T111, Houston,  
TX 77030, USA  
david@eaglemanlab.net

### Specialty section:

This article was submitted to  
Perception Science,  
a section of the journal  
Frontiers in Psychology

**Received:** 26 March 2015

**Accepted:** 08 July 2015

**Published:** 12 August 2015

### Citation:

Cai MB and Eagleman DM (2015)  
Duration estimates within a modality  
are integrated sub-optimally.  
Front. Psychol. 6:1041.  
doi: 10.3389/fpsyg.2015.01041

Perceived duration can be influenced by various properties of sensory stimuli. For example, visual stimuli of higher temporal frequency are perceived to last longer than those of lower temporal frequency. How does the brain form a representation of duration when each of two simultaneously presented stimuli influences perceived duration in different way? To answer this question, we investigated the perceived duration of a pair of dynamic visual stimuli of different temporal frequencies in comparison to that of a single visual stimulus of either low or high temporal frequency. We found that the duration representation of simultaneously occurring visual stimuli is best described by weighting the estimates of duration based on each individual stimulus. However, the weighting performance deviates from the prediction of statistically optimal integration. In addition, we provided a Bayesian account to explain a difference in the apparent sensitivity of the psychometric curves introduced by the order in which the two stimuli are displayed in a two-alternative forced-choice task.

**Keywords:** duration perception, cue integration, memory decay, Bayesian inference, temporal frequency, time order error, just noticeable difference

## Introduction

Estimating how long an event lasts is a perceptual capacity that we utilize in daily life. For example, we distinguish words with similar sounds, such as “sheep” and “ship,” based on the duration of a syllable; a salesman can infer a customer’s interest by how long the customer gazes on each item; we judge internet speed based on the time it takes to load a webpage; various electric devices signal different messages to us by the duration of a beep or flash. However, the mechanisms by which the brain estimates a duration is still unclear (For an non-exhaustive list of recent reviews on duration perception, see Eagleman, 2008; Ivry and Schlerf, 2008; Grondin, 2010; Merchant et al., 2013). A traditional view of duration perception is that the brain possesses a dedicated “internal clock” (Treisman, 1963; Gibbon, 1977). In this view, duration perception is less dependent on low-level sensory processing. However, recent psychophysical studies have revealed that perceived duration can, in fact, be influenced by various properties of a visual stimulus, such as temporal frequency or speed of motion (Brown, 1995; Kanai et al., 2006; Kaneko and Murakami, 2009; Tomassini et al., 2011; Kline and Reed, 2013), change of speed (Carrozzo and Lacquaniti, 2012), numerosity (Long and Beaton, 1981; Xuan et al., 2007), contrast (Long and Beaton, 1980; Xuan et al., 2007), spatial frequency (Aaen-Stockdale et al., 2011), and looming (van Wassenhove et al., 2008). The fact that duration perception is influenced by so many low-level sensory features suggests that the details of a sensory stimulus contribute to its perceived duration. Perceived duration is not only influenced by the property of sensory stimuli, but also by the history of stimuli: a repeated stimulus appears briefer than a novel stimulus (Tse et al., 2004; Pariyadath and Eagleman, 2007; Schindel et al., 2011; Birngruber et al., 2014). This phenomenon

has been suggested to reflect a link between neural response amplitude and perceived duration (Pariyadath and Eagleman, 2007; Eagleman and Pariyadath, 2009). In addition, it was found that after adaptation to a fast drifting visual stimulus, a slow drifting visual stimulus is perceived as being of shorter duration when it appears at the adapted visual field, but not at other locations (Johnston et al., 2006; Ayhan et al., 2009, 2011; Bruno et al., 2010). The latter example not only highlights the involvement of low-level sensory processing in duration perception, but also demonstrates that stimuli in different parts of the visual field can provide different evidence of duration.

The finding that perceived duration can be biased by the sensory features of stimuli creates a puzzle. Even if visual objects at different locations last for the same physical duration, they each can bias perceived duration in different directions due to their sensory features. How does the brain form a representation of duration based on the duration estimates from different visual objects?

One possibility, as an extension of the hypothesis that perceived duration is based on neural response amplitude (Eagleman and Pariyadath, 2009), is that the perceived duration may be based on the sum of the total neural response to all the stimuli. An alternative hypothesis is that an estimate of duration is formed based on each stimulus and the brain integrates these estimates by a weighted average. A stronger statement of this hypothesis is that the integration may be statistically optimal (Ahrens and Sahani, 2011). A third hypothesis is that the brain may form a duration representation based on only one of the stimuli, with certain probability. A fourth hypothesis is that the brain may only rely on the stimulus type that provides more reliable (less variable) estimate of duration across trials. Lastly, it is possible that the brain may generate a representation of duration based on each stimulus and keep all the representations. In this last framework, the brain may have flexibility to choose which representation to use depending on the task.

Closely related to the question asked in this study, Ayhan et al. (2012) investigated whether human observers can average the durations of multiple objects. They flashed multiple images of different durations with asynchronous onsets and asked participants to make judgments with regards to the average duration of those images. The precision of the duration judgment was found to be worse when judging the average duration of multiple images than when judging the duration of a single image. The authors suggested that this reflects an inability to aggregate duration information from multiple items (Ayhan et al., 2012). While this may be the case when the stimuli have asynchronous onsets and offsets, there has been no study investigating whether and how human observers combine duration information from multiple objects which appear and disappear synchronously. To study the combination of duration information without introducing asynchrony between stimuli, we utilize the illusion that the temporal frequency of a visual stimulus biases perceived duration to create conflicting estimates of duration. In Experiment 1, we confirm this illusion by a two-alternative forced choice task. In Experiment 2, we qualitatively test the predictions of each of the above hypotheses to focus our attention on a few most plausible candidate models. In

Experiment 3, we quantitatively compare these candidate models based on the trial-by-trial cross-validated log-likelihood of the models.

## Participants and Methods

The experiments were approved by the Institutional Review Board of Baylor College of Medicine.

### Participants

Except for the first author, participants were all naïve to the purpose of the study. Participants provided informed consent and received compensation. Nineteen participants (8 males, 11 females. Age  $27 \pm 7$ ) took part in Experiment 1. Twenty-one participants (13 males, 8 females. Age  $29 \pm 7$ ) took part in Experiment 2. Twenty participants (6 males, 14 females. Age  $27 \pm 6$ ) took part in Experiment 3.

### Apparatus

Experiment stimuli were displayed on a CRT monitor (Viewsonic G225f) with a screen resolution of  $1024 \times 768$  pixels and a refresh rate of 100 Hz, driven by a Dell Precision T3400 workstation running Windows XP. There was no other light source other than the monitor in the experimental room. Participants sat at a distance of approximately 60 cm from the display. Each participant wore a pair of earplugs with approximately 33 dB noise reduction to prevent distraction.

### Stimuli

Stimuli were presented using Psychtoolbox 3 (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007) for Matlab. Stimuli consisted of one or two drifting Gabor patches with spatial frequency of 0.28 cycle/degree (estimated at 60 cm viewing distance). The standard deviation of the 2-dimensional Gaussian envelop of each Gabor patch was  $0.90^\circ$ . The starting phase of each Gabor patch was independently sampled from a uniform distribution over the range of  $0-2\pi$ . The peak luminance of the Gabor patch was  $36.0 \text{ cd/m}^2$ . Stimuli were presented over gray background of mid-luminance. Each Gabor patch was displayed at a distance of  $5.4^\circ$  visual angle away from the fixation point. The fixation point was at the center of the screen, indicated by a white cross spanning a visual angle of  $0.6^\circ$ . Through the time course of each stimulus, the sinusoidal component of each Gabor patch drifted in a direction independently sampled from a uniform distribution over the range of  $0-360^\circ$ . The speed of their drifting was such that the luminance of any pixel of the Gabor patch was modulated by a sinusoidal time signal of either 1 Hz (for the low temporal frequency stimulus) or 6 Hz (for the high temporal frequency stimulus). At the onset of each stimulus, the contrast of the Gabor patch ramped up linearly from zero to maximum in 40 ms. At the offset, it ramped down in 40 ms. This ramping of the contrast was to minimize potential arousal introduced by abrupt onsets of stimuli.

Whenever two Gabor patches were displayed simultaneously, the centers of the two Gabor patches were on opposite sides from the fixation point, both on an invisible line that passed through the fixation point. In any trial, the orientation of the invisible



line passing through the fixation point and the Gabor patch(es) in the first epoch was randomly sampled from a uniform distribution over  $0-2\pi$ . The invisible line passing through the fixation point and the Gabor patch(es) in the second epoch was always orthogonal to the invisible line in the first epoch. This design was to minimize the effect of adaption due to presenting consecutive stimuli at the same location (Johnston et al., 2006).

## Experiment Procedures

On each trial, a participant watched two groups of drifting Gabor patterns on the screen one after another and judged whether the duration of the second group was longer or shorter than that of the first group. Each group was composed of either a single Gabor patch drifting at 1 Hz (we denote this by L), or a single Gabor patch drifting at 6 Hz (we denote this by H), or a pair of Gabor patches, one at 1 Hz and the other at 6 Hz (we denote this by HL). In an HL stimulus, the two Gabor patches had the same onset time and offset time. The directions in which they drifted were randomly chosen and independent from each other. If a participant asked which one patch of the HL stimulus they should judge, he/she was instructed that since the patches appeared and disappeared synchronously, he/she should judge the duration in which both of them stay on the screen.

The structure of each trial was as follows. A trial started by a fixation cross appearing in the center of the screen. After a duration sampled from a uniform distribution over the range of 600–1000 ms, the first group of Gabor patch(es) appeared. 500–700 ms after the offset of the first group of Gabor patch(es), the second group appeared. 300–600 ms after the offset of the second group, the fixation cross disappeared and the participants were allowed to make response. They indicated the duration of the second group as lasting longer by pressing the right arrow key, or indicated it as lasting shorter by pressing the left arrow key. No feedback was provided. 1000–2000 ms after they made a response, the next trial started.

On any trial of an experiment, one group of Gabor patches lasted for 600 ms. We denote this stimulus of fixed duration by reference stimulus. The other group lasted for duration of one of 26 values between 100 and 1100 ms, equally spaced by steps of 40 ms. We denote this stimulus by comparison stimulus. For each of these 26 values, the number of its incidence was approximately proportional to the probability density of a Gaussian distribution with a mean of 600 ms and a standard deviation of 300 ms at that duration, rounded to the nearest integer. Thus, over the course of an experiment, the distribution of the duration of comparison stimuli approximates a truncated Gaussian distribution.

## Experiment 1

There were two conditions in the experiment. In one condition, the reference stimulus was H and the comparison stimulus was L (denoted by LvsH). In the other condition, the reference was L and the comparison was H (denoted by HvsL). On half of the trials of each condition, the reference stimulus appeared before the comparison stimulus. On the other half of the trials, the comparison stimulus appeared before the reference stimulus. Each condition had 180 trials, including both orders of display. For each order of display in each condition, the comparison stimuli of 100, 140, 180, ..., and 1100 ms occurred for 1, 2, 2, 2, 3,

3, 4, 4, 4, 5, 5, 5, 5, 5, 5, 5, 4, 4, 4, 3, 3, 2, 2, 2, and 1 times. These numbers of incidences were generated to approximate a Gaussian distribution described above. Trials corresponding to different conditions, orders and comparison durations were randomly interleaved in a session. There was no signal to indicate to the participants which condition a trial belonged to.

## Experiment 2

On all trials, the reference stimulus was an HL stimulus. The comparison stimulus was an L, H, or HL stimulus. The reference stimulus was always presented before the comparison stimulus. Each condition had 148 trials. In each condition, the comparison stimuli of 100, 140, 180, ..., and 1100 ms occurred for 2, 2, 4, 4, 4, 6, 6, 6, 8, 8, 8, 8, 8, 8, 8, 8, 6, 6, 6, 4, 4, 2, and 2 times. The trials of the three conditions were randomly interleaved.

## Experiment 3

There were seven conditions in the experiment. In two conditions, the reference stimulus was H; the comparison stimulus was H or L, respectively. In two other conditions, the reference stimulus was L; the comparison stimulus was H or L, respectively. In the other three conditions, the reference stimulus was HL; the comparison stimulus was H, L, or HL, respectively. On half of the trials of each condition the reference stimulus was presented before the comparison stimulus. On the other half of the trials, the comparison stimulus was presented before the reference stimulus. Each condition had 228 trials. Each participant completed three sessions of experiment. For each order of display in each condition, the comparison stimuli of 100, 140, 180, ..., and 1100 ms occurred for 3, 3, 3, 3, 3, 3, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 3, 3, 3, 3, 3, 3, and 3 times in total over all sessions. Trials corresponding to different conditions, orders and durations of comparison stimuli were randomly interleaved in a session. The number of trials corresponding to each condition, order and duration of comparison stimulus was equal across sessions.

# Results

## Experiment 1

It has been found that visual stimuli of higher temporal frequency or faster speed are perceived as lasting for longer than those of lower temporal frequency or slower speed (Kanai et al., 2006; Kaneko and Murakami, 2009). Our goal in Experiment 1 is to confirm this finding. In the previous literature, the overestimation of duration was measured by a reproduction task: after watching a stimulus, participants pressed a button for as long as they believed the stimulus had lasted. The variance of the reproduced duration in such a task is contributed to by the variance of participants' perceived duration and the noise in their motor timing. To avoid the latter, we used a two-alternative forced choice task, in which participants watched two consecutive stimuli and judged which lasted longer. This offers a more accurate estimation of the difference in perceived durations between stimuli of high and low temporal frequencies.

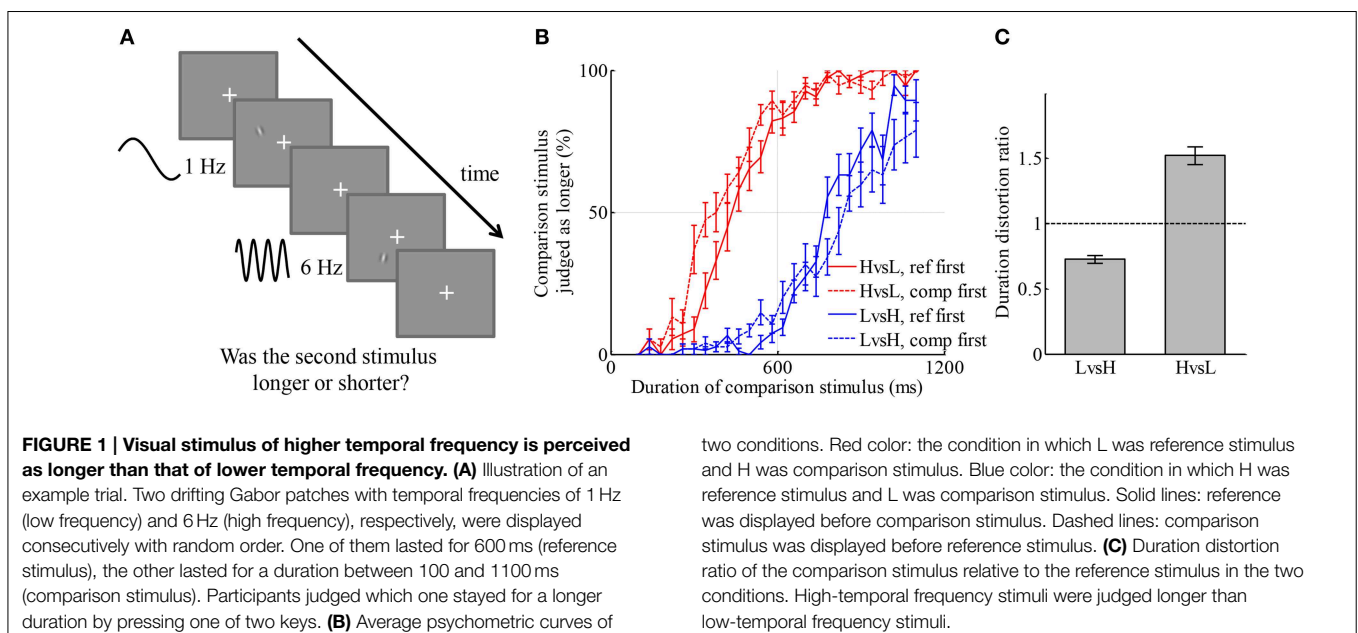
The stimuli of an example trial are shown in **Figure 1A**. Each stimulus was a supra-threshold Gabor patch. Each pixel of the Gabor patch was modulated by a sinusoidal time series

of either 1 Hz (we denote this low frequency by L) or 6 Hz (we denote this high frequency by H). Thus, the Gabor patch appeared as a grating that drifted behind a static 2-dimensional Gaussian aperture. The first Gabor patch appeared at a random location with fixed distance from the center of the screen (fixation point). The second Gabor patch appeared at the same distance from fixation but either 90° clockwise or counterclockwise from the first Gabor patch. On any trial, one of the stimuli lasted for 600 ms (we denote this as reference stimulus), and the other lasted for one of 26 durations equally spaced between 100 and 1100 ms (we denote this as comparison stimulus). The distribution of the duration of the comparison stimulus approximated a truncated Gaussian distribution with mean of 600 ms and standard deviation of 300 ms. On half of the trials, the comparison stimulus was H and the reference stimulus was L (HvsL condition). On the other half of the trials, the comparison stimulus was L and the reference stimulus was H (LvsH condition). On half the trials of each condition, the reference stimulus appeared before the comparison stimulus. On the other half, it appeared after. Participants reported whether the second stimulus lasted longer or shorter than the first stimulus.

The participant-averaged psychometric curves are displayed in **Figure 1B**. A leftward shift of a curve from centering at 600 ms indicates that the duration of the comparison stimulus was overestimated relative to the reference stimulus, and vice versa for a rightward shift. There was a slight discrepancy between the curves corresponding to different orders of display, namely, that curves deviated more from the reference duration and were shallower when the comparison stimulus was presented first. This type of discrepancy was also found in many other studies of perceptual judgments (Nachmias, 2006; Lapid et al., 2008; Bruno et al., 2010, 2012; Ahrens and Sahani, 2011). We will investigate the source of such discrepancy in Experiment 3, together with quantitatively comparing models of the representation of

duration for simultaneously presented H and L stimuli. For simplicity, trials of different orders of display but belonging to the same condition were aggregated in the analysis. We fitted each participant's responses in each condition by a curve of Gaussian cumulative distribution on the logarithmic scale of duration, with an additional term capturing lapse rate, the chance that a participant had not paid attention to the stimuli (Wichmann and Hill, 2001). The ratio of the perceived duration of comparison stimuli to that of reference stimuli in each condition was calculated based on the exponential of the shift of the psychometric curve in the logarithmic scale. We denote this ratio by the duration distortion ratio (DDR, **Figure 1C**). In the LvsH condition, the duration of the L stimulus was judged as  $27.3 \pm 3.0\%$  (mean  $\pm$  s.e.m, the same through this paper unless otherwise stated) shorter than the H stimulus; the DDR was significantly smaller than 1 [ $t_{(18)} = -9.10, p < 0.001$ ]. In the HvsL condition, the duration of the H stimulus was judged as  $52.1 \pm 6.8\%$  longer than the L stimulus; the DDR was significantly larger than 1 [ $t_{(18)} = 7.67, p < 0.001$ ]. The standard deviations of the fitted Gaussian cumulative distribution functions represent participants' sensitivity in discriminating duration in the two conditions, termed as just noticeable difference (JND). The JND was  $0.27 \pm 0.03$  on the logarithmic scale of duration in the LvsH condition, and  $0.35 \pm 0.03$  in the HvsL condition. They were significantly different [ $t_{(18)} = -3.99, p < 0.001$ ]. The JND in logarithmic scale has similar meaning to Weber's ratio. When psychometric curves were fitted without applying logarithmic transformation of duration, the conclusions about DDR and Weber's ratio stayed the same. The absolute value of the DDR is very different between LvsH and HvsL conditions. This may indicate that the distortion in perceived duration caused by the temporal frequency is multiplicative instead of additive.

Experiment 1 confirms the previous finding that the perceived duration of visual stimulus is biased by its temporal frequency



or speed. This leads to our main question: how do we perceive duration if two stimuli are presented simultaneously, one of which moves faster and the other slower. In Experiment 2, we test several hypotheses.

## Experiment 2

This experiment examined the perceived duration of two stimuli appearing simultaneously at different locations, one of low temporal frequency (L) and one of high temporal frequency (H). We denote such stimuli by HL. The H and L elements of it appear and disappear at the same time. This provides a clue that they should correspond to the same period of duration. However, following the observation in Experiment 1, the H and L elements of HL each should cause conflicting biases on the respective duration estimates, with H indicating a longer duration and L indicating a shorter duration. How does the brain form a representation of duration for the joint stimulus?

We consider five possibilities:

### Global Summing Hypothesis

It is noticeable that neural response amplitude in visual cortex also increases with temporal frequency in the range that was tested in Kanai et al.'s experiments (Singh et al., 2000). The bias in perceived duration caused by the temporal frequency or speed of visual stimuli may be explained by assuming that perceived duration is based on the neural response amplitude to the stimulus (Eagleman and Pariyadath, 2009). It may also be explained by assuming that duration perception is based on the number of changes observed (Brown, 1995; Kanai et al., 2006). As possible extensions of both of these hypotheses, we may assume that the perceived duration of multiple elements is based on either the total neural responses to all the stimulus elements or the total number of changes in all stimulus elements. We denote such hypotheses by “global summing.” Both of them predict that HL should be perceived as lasting longer than both H and L.

### Weighting Hypothesis

The perceived duration of HL may be formed by a weighted average of each estimate of duration based on one of its elements. We denote by  $x_H$  the estimate of duration based on an H stimulus lasting for a physical duration of  $t$ , and denote by  $x_L$  the one based on an L stimulus lasting the same duration.  $x_H$  and  $x_L$  both vary across trials. We assume that their variations are independent and both follow Gaussian distributions:

$$x_H \sim N(t + b_H, \sigma_H) \quad (1)$$

$$x_L \sim N(t + b_L, \sigma_L) \quad (2)$$

$b_H$  and  $b_L$  represent the bias of perceived duration introduced by their temporal frequencies.  $\sigma_H$  and  $\sigma_L$  represent the standard deviation of the distribution of  $x_H$  and  $x_L$ . For simplicity, we assume that a point estimation of the duration of stimulus HL is formed by weighting  $x_H$  and  $x_L$ :

$$x_{HL} = w_H x_H + (1 - w_H) x_L \quad (3)$$

where the weight  $w_H$  is a parameter of each participant, in the range of [0, 1]. The distribution of  $x_{HL}$  would follow:

$$x_{HL} \sim N(t + w_H b_H + (1 - w_H) b_L, \sqrt{w_H^2 \sigma_H^2 + (1 - w_H)^2 \sigma_L^2}) \quad (4)$$

For any weight  $w_H$ , this predicts that on average HL is perceived equal to or shorter than H, and equal to or longer than L. The equality is only reached if  $w_H$  is 0 or 1, meaning one of the elements is neglected. It also predicts that the standard deviation of the perceived duration of HL is equal to or smaller than the larger one of those of H and L (namely,  $\sigma_{HL} \leq \max\{\sigma_H, \sigma_L\}$ ). The equality is only reached when the duration estimation is only based on the more variable estimation between  $x_H$  and  $x_L$ , i.e., when  $w_H = 1$  and  $\sigma_H \geq \sigma_L$ , or when  $w_H = 0$  and  $\sigma_H \leq \sigma_L$ .

The statistically optimal way to weight sensory evidence is by setting the weight of each duration estimation inversely proportional to the variance of that estimation (Jacobs, 1999; Knill and Pouget, 2004). We denote the hypothesis that the weighting follows this rule as the “optimal integration” hypothesis, as a stronger version of the “weighting” hypothesis. Based on this hypothesis, we expect the perceived duration of HL to be less variable than that of each stimulus element H and L:

$$\sigma_{HL} = \sqrt{\frac{\sigma_H^2 \sigma_L^2}{\sigma_H^2 + \sigma_L^2}} < \min\{\sigma_H, \sigma_L\} \quad (5)$$

### Selection Hypothesis

Instead of weighting the estimates based on the two stimulus elements, the brain may estimate the duration based on only one of the two elements. On some trials the perceived duration may be based on the H element and on other trials it is based on the L element. The element selected to form duration representation on a trial may be the one which more attention is paid to. Assuming a participant has a probability  $c_H$  to rely on the H element to estimate duration, we have

$$x_{HL} = \begin{cases} x_H, & \text{with probability } c_H \\ x_L, & \text{with probability } (1 - c_H) \end{cases} \quad (6)$$

With the same notation as we used above, the mean of  $x_{HL}$  across trials would be

$$t + c_H b_H + (1 - c_H) b_L \quad (7)$$

and the standard deviation of  $x_{HL}$  across trials would be

$$\sqrt{c_H \sigma_H^2 + (1 - c_H) \sigma_L^2 + c_H (1 - c_H) (b_H - b_L)^2} \quad (8)$$

This predicts that the average of the perceived duration of HL across trials is also equal to or shorter than that of H, and equal to or longer than that of L. Equality is only reached if  $c_H$  is equal to 0 or 1. As opposed to the “weighting” hypothesis, it predicts that the standard deviation of the perceived duration of HL across trials is equal or larger than the smaller one of those of H and L (namely,  $\sigma_{HL} \geq \min\{\sigma_H, \sigma_L\}$ ). The equality is only reached when the duration representation is always based on the stimulus type which gives rise to a smaller variance of duration estimation, i.e., when  $c_H = 1$  and  $\sigma_H < \sigma_L$ , or when  $c_H = 0$  and  $\sigma_H > \sigma_L$ .

## Reliable Stimulus Hypothesis

The brain might only rely on one of the stimulus types across all the trials, and the stimulus type it relies on may be the one that in general gives rise to more reliable estimation of duration. Under this hypothesis, if a participant estimates the duration of H with less variability than estimating the duration of L, the participant may always estimate the duration of HL based on the H element. If the participant estimates the duration of L with less variability, he/she may always rely on the L element to estimate the duration of HL. This hypothesis also predicts that  $\sigma_{HL} \leq \max\{\sigma_H, \sigma_L\}$ . The average perceived duration of HL may be shorter than that of H and longer than that of L across participants, if not all participants estimate a same type of stimulus between H and L more reliably than the other. However, for those who have more reliable estimates of duration based on H, the perceived duration of HL should be on average equal to that of H. And similarly for those who have more reliable estimates of duration based on L.

## Multiple Representations Hypothesis

Instead of forming a single representation of duration as assumed by the above hypotheses, the brain might keep multiple representations of duration, each based on one of the two simultaneously presented stimuli. When asked to compare the duration of HL with the duration of a single stimulus, the brain might use one of the two representations formed during HL that is based on the stimulus element that is most similar to the single stimulus to be compared. For example, when viewing HL, the brain might keep one duration representation based on H and one based on L. When asked to compare the duration of HL with the duration of H, the brain might compare the representation based on the H element of HL with the duration representation of the single H stimulus. In this case, H should be judged to be of the same duration as HL on average. Similarly, L should also be judged equally long as HL. In other words, under this hypothesis, when the reference stimulus is HL and the comparison stimulus is H or L, the DDRs of H and L relative to HL should be equal.

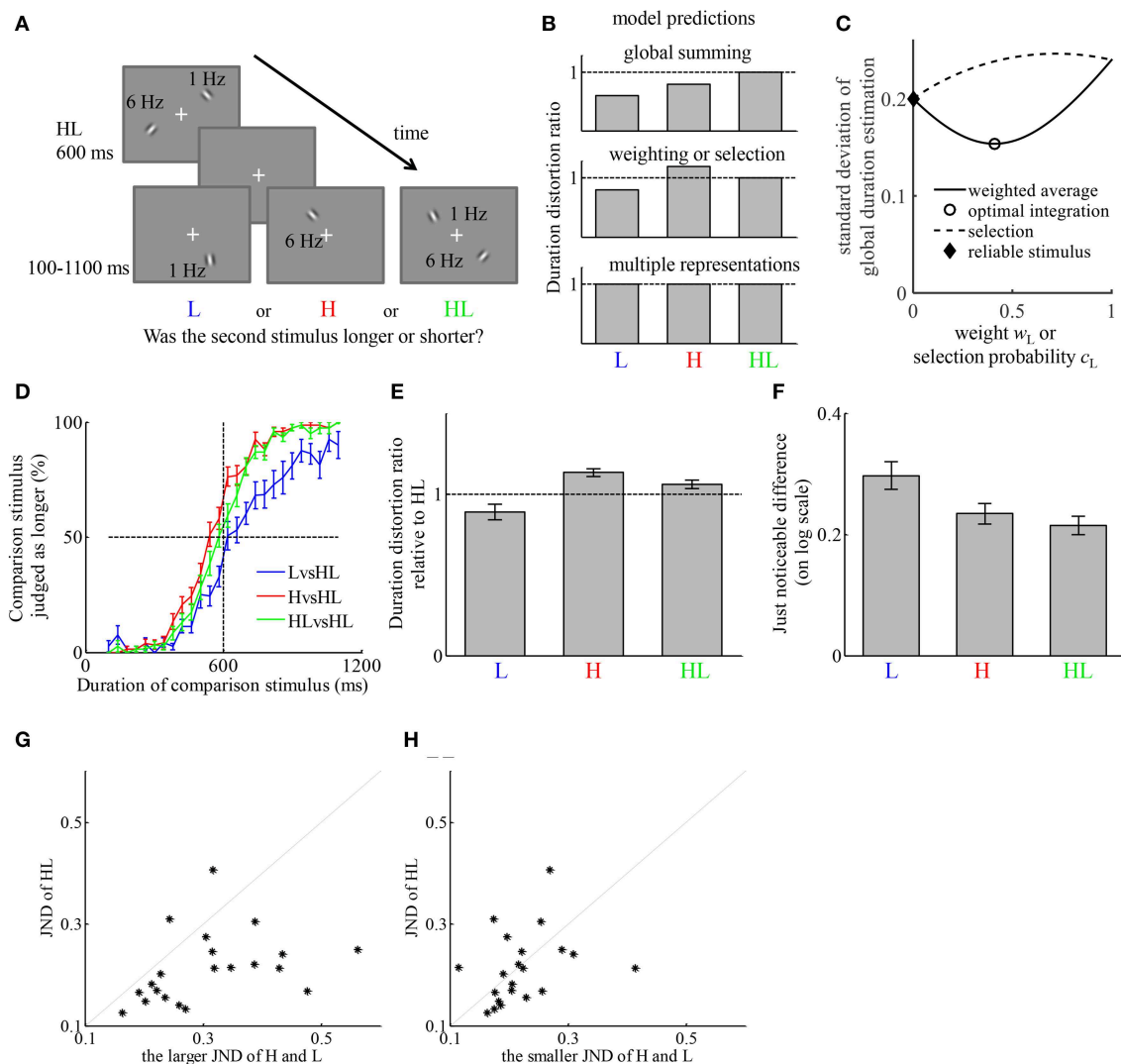
To test the above predictions, we asked participants to compare the duration of H, L, or HL against the duration of HL. Example trials are shown in **Figure 2A**. On each trial, the reference stimulus was always presented before the comparison stimulus. The reference stimuli were all of HL type. There were three conditions distinguished by the types of comparison stimuli. In 1/3 of the trials, the comparison stimuli were L (LvsHL condition). In 1/3, the comparison stimuli were H (HvsHL condition). In the other 1/3, the comparison stimuli were HL (HLvsHL condition). Trials of the three conditions were randomly interleaved. Participants judged whether the duration of the second stimulus was longer or shorter than that of the first on each trial.

We tested the predictions of each of the models by comparing the DDRs between conditions. Each of the hypotheses generates prediction about the relation between the average perceived duration of HL and those of H and L. **Figure 2B** provides a qualitative illustration of their differences. The “weighting” and “selection” hypotheses generate the same qualitative prediction about the average perceived duration of HL. The “reliable stimulus” hypothesis may generate similar prediction as these

two as long as there is individual difference regarding which of H and L is estimated with less variability. They are further distinguished by their qualitative predictions of  $\sigma_{HL}$ , the standard deviation of perceived duration of HL. Without losing generality, by fixing the values of  $\sigma_H$ ,  $\sigma_L$  and  $b_H - b_L$ , **Figure 2C** illustrates how  $\sigma_{HL}$  varies as a function of  $w_H$  or  $c_H$ , which are both free parameters of each participant. The “weighting” hypothesis predicts  $\sigma_{HL} \leq \max\{\sigma_H, \sigma_L\}$  while the “selection” hypothesis predicts  $\sigma_{HL} \geq \min\{\sigma_H, \sigma_L\}$ . Under the “optimal integration” hypothesis, a stronger version of the “weighting” hypothesis, we have  $\sigma_{HL} \leq \min\{\sigma_H, \sigma_L\}$ . The “reliable stimulus” hypothesis predicts  $\sigma_{HL} \leq \max\{\sigma_H, \sigma_L\}$ . The predictions about the average perceived duration of HL are tested by comparing the DDRs of each stimulus type relative to HL. Although the standard deviations of perceived duration of each stimulus type cannot be directly measured, they have monotonic relation with the JNDs in each condition. Therefore, the predictions about the standard deviations of perceived duration are tested by comparing the JNDs between conditions.

The participant-averaged psychometric curves are displayed in **Figure 2D**. We fitted each participant’s responses similarly as in Experiment 1. The DDRs of the three conditions are displayed in **Figure 2E**. In the LvsHL condition, the duration of the L stimulus was judged as  $11.0 \pm 4.8\%$  shorter than HL stimulus. In the HvsHL condition, the duration of the H stimulus was judged as  $13.3 \pm 2.5\%$  longer than the HL stimulus. In the HLvsHL condition, the duration of HL as comparison stimulus was judged as  $5.9 \pm 2.7\%$  longer than the HL as reference stimulus. A repeated measures ANOVA revealed a significant difference in DDR between the three conditions [ $F_{(2, 40)} = 11.81, p < 0.001$ ]. *Post-hoc* paired *t*-tests between each two conditions revealed a significant difference between the LvsHL and HvsHL conditions [ $t_{(20)} = -4.21, p < 0.001$ ], a significant difference between the LvsHL and HLvsHL conditions [ $t_{(20)} = -2.66, p = 0.015$ ] and a significant difference between the HvsHL and HLvsHL conditions [ $t_{(20)} = 3.33, p = 0.003$ ], all of which passed the Holm-Bonferroni multiple comparison criterion (Holm, 1979). The DDR in HvsHL condition was significantly larger than 1 (*t*-test,  $p < 0.001$ ). The DDRs in the LvsHL was on average smaller than 1, but the difference was not significant after correcting for multiple comparison ( $p = 0.03$ , Holm-Bonferroni criterion). The DDR in the HLvsHL condition was also not significantly different from 1 ( $p = 0.04$ , Holm-Bonferroni criterion). The JNDs of the three conditions are shown in **Figure 2F**. Because the psychometric functions were fitted after logarithmic transformation of the duration, their units are also in the logarithmic scale. A repeated measures ANOVA revealed significant difference in JNDs between the three conditions [ $F_{(2, 40)} = 7.48, p = 0.002$ ]. *Post-hoc* paired *t*-test between each pair of conditions revealed a significant difference between LvsHL and HvsHL conditions [ $t_{(20)} = 2.81, p = 0.011$ ], a significant difference between the LvsHL and HLvsHL conditions [ $t_{(20)} = 3.57, p = 0.002$ ], but no significant difference between the HvsHL and HLvsHL conditions [ $t_{(20)} = -0.02, p = 0.31$ ]. The JND in the HLvsHL condition was significantly smaller than the maximum of those in the other two conditions [ $t_{(20)} = -4.23, p < 0.001$ ], (**Figure 2G**) but not significantly





**FIGURE 2 | The representation of duration of simultaneously presented high- and low-temporal frequency stimuli (HL) can be described by a weighted average of the estimates of duration based on the high-temporal frequency stimulus element (H) and low-temporal frequency stimulus element (L).** (A) Example of the stimuli on a trial. Participants first viewed an HL stimulus lasting for 600 ms, then viewed one of three types of stimuli, H, L, or HL, with variable duration between 100 and 1100 ms. Participants judged which one lasted longer.

(B) The qualitative relation between the duration distortion ratios of the comparison stimulus relative to the reference stimulus, predicted by four hypotheses of how the representation of the duration of HL is formed. The "reliable stimulus" hypothesis may generate the same prediction as

"weighting" and "selection" hypotheses if not all participants estimate the same type of stimulus more reliably. (C) Illustration of the different predictions of the standard deviation of perceived duration of HL in comparison to that of H and L of the "weighting," "optimal integration," "selection," and "reliable stimulus" hypotheses. The figure is generated by assuming  $\sigma_H = 0.2$ ,  $\sigma_L = 0.24$ , and  $b_H - b_L = 0.2$ . (D) Average psychometric curves of the three conditions. (E) Average duration distortion ratio of the three conditions. (F) Average just noticeable difference (JND) of the three conditions. (G) Comparison between the JND in the HLvsHL condition and the larger JND of the other two conditions. Each dot corresponds to one participant. (H) Comparison between the JND in the HLvsHL condition and the smaller JND of the other two conditions.

different from the minimum of those in the other conditions [ $t_{(20)} = -0.40, p = 0.69$ ] (Figure 2H).

The finding that HL was judged shorter than H argues against the "global summing" hypothesis. The "multiple representations" hypothesis is also ruled out because H and L was judged differently relative to HL stimulus. The pattern of DDRs among conditions of this experiment is consistent with both the "weighting" and "selection" hypotheses. The key difference of their predictions is with the standard deviation

of the duration estimation of HL compared to those of H and L. JND indirectly reflects the standard deviation. The finding that JND in HLvsHL condition was smaller than the maximum of the JNDs in the other conditions supports the "weighting" and "reliable stimulus" hypotheses. The finding that it was not significantly different from the minimum of the JNDs in the other conditions does not provide support to the "selection" hypothesis or the "optimal integration" hypothesis. If the "reliable stimulus" hypothesis is true, then the participants who

estimate the duration of H with less variability than L should have no difference in DDR between the HLvsHL and HvsHL conditions; the participants who estimate the duration of L with less variability should have no difference in DDR between the HLvsHL and LvsHL condition. Because the JND is smaller in HvsHL condition for majority of the participants (16 out of 21), we test the former prediction in these participants. The DDR was on average smaller in the HLvsHL condition ( $7.3 \pm 3.2\%$ ) than in the HvsHL condition ( $12.5 \pm 2.6\%$ ). The difference was marginally significant with  $p = 0.054$ .

We also note that the DDR in the HLvsHL condition was larger than 1, although the significance level did not pass our multiple comparison threshold. This may be due to participants' response bias or their prior belief about the relation between the first and second stimuli. However, such factors should equally impact all conditions. They do not influence our conclusions because the conclusions are based on comparisons between conditions. When psychometric curves were fitted without taking a logarithmic transform of duration, all conclusions remained the same except that the JNDs in LvsHL and HvsHL were not significantly different ( $p = 0.14$ ), which was not crucial for testing the model predictions.

Therefore, the result of Experiment 2 provided qualitative evidence that the perceived duration of two dynamic stimuli is more likely formed by weighting the estimates of duration based on each individual stimulus, although we cannot entirely rule out the "reliable stimulus" hypothesis.

### Experiment 3

Experiment 2 ruled out the "global summing" and "multiple representations" hypotheses, provided qualitative support to the "weighting" hypothesis, but could not rule out the "reliable stimulus" hypothesis. The predictions of the "selection" and "optimal integration" hypotheses were not supported by the data, but they were also not entirely ruled out. In order to formally compare the "weighting" hypothesis, the "optimal integration" hypothesis, the "selection" hypothesis and the "reliable stimulus" hypothesis, one needs to explicitly model the decision process of each trial, predict the probability that a participant makes each judgment, and calculate the likelihood of each model. The probability that one stimulus is judged longer than another depends on both the mean and standard deviation of the perceived duration of the two stimuli over repetition of trials. As shown in Equations (4), (6), and (7), under each hypothesis, the mean and standard deviation of perceived duration of HL depends on those of the perceived durations of both H and L. Experiment 3 additionally included conditions in which the two stimuli on a trial were H and H, L and L, and H and L. These conditions constrained the fitting of parameters corresponding to the means and standard deviations of perceived duration of H and L, namely  $b_H$ ,  $b_L$ ,  $\sigma_H$ , and  $\sigma_L$ . In Experiment 1 we noticed a discrepancy in psychometric curves corresponding to different orders in which reference and comparison stimuli were displayed. To investigate the source of this discrepancy, trials of both orders of display were included for each condition in Experiment 3.

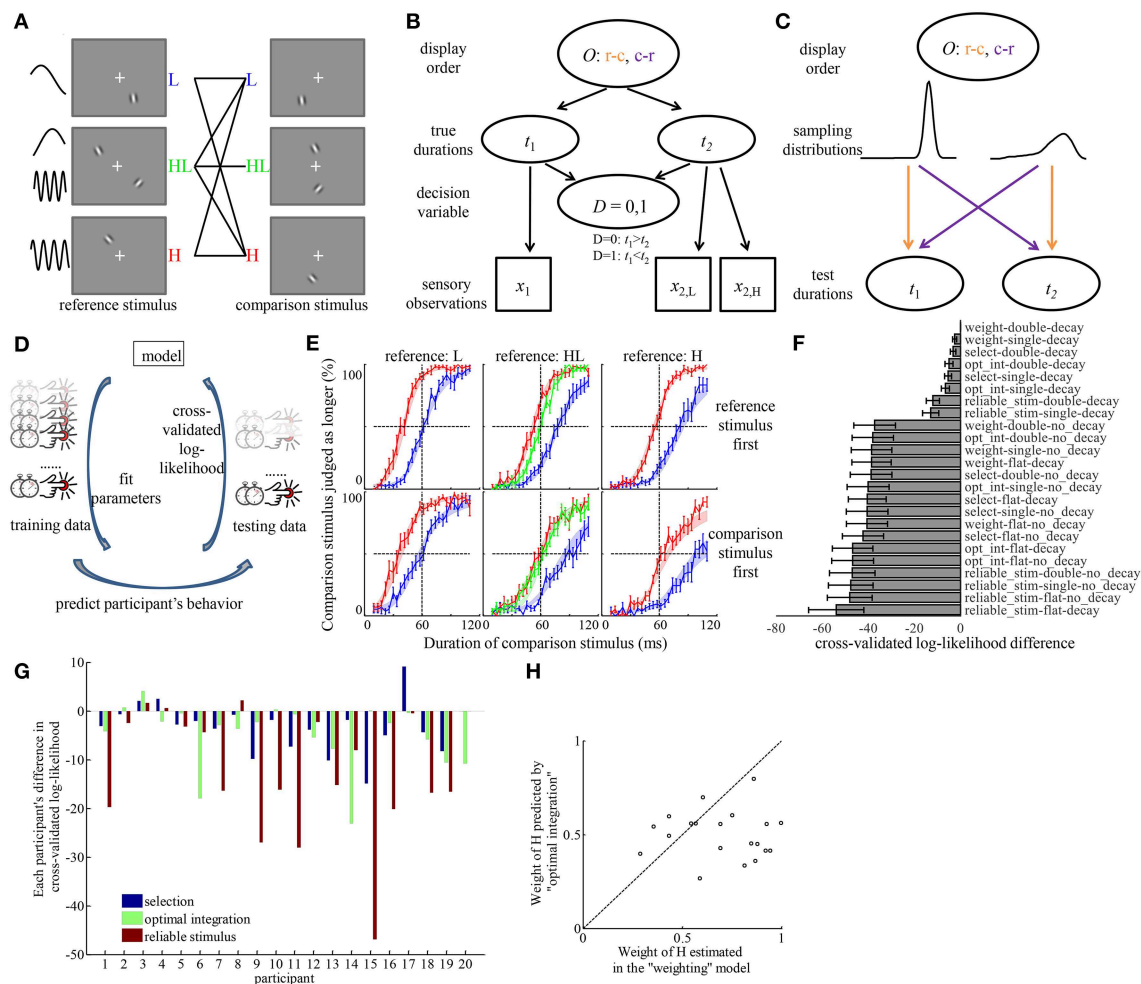
The timing structure of a trial in Experiment 3 was the same as in Experiment 1. There were seven conditions, defined by their reference and comparison stimuli. These conditions are illustrated in **Figure 3A**. The participant-averaged psychometric curves of each condition and each order of display are shown in **Figure 3E**. Similarly to Experiment 1, a discrepancy existed between the orders of displaying the reference and comparison stimuli. In general, psychometric curves were steeper and closer to the center of the range of duration when the reference stimulus was displayed first.

In order to understand the process of forming the representation of duration of HL and the discrepancy in judgments due to the order of display, we constructed models based on different hypotheses concerning three factors (van den Berg et al., 2014), and compared the log-likelihood of each model by cross-validating it within data of each participant. The details of the model comparison approach are described in Data Analysis and Modeling. Here we briefly list the major steps.

We consider the generative model of the sensory measurements of duration by the brain as in **Figure 3B**. The two durations to be compared on any trial were sampled from two distributions, one corresponding to the reference stimulus, and one corresponding to the comparison stimulus, as illustrated in **Figures 3B,C**. The order in which they were displayed was random from trial to trial. The true durations should be unknown to the brain. The brain only has sensory measurements of duration based on each of H or L stimulus, or each element of HL stimulus, which are noisy and biased by the temporal frequencies. We assume that the brain infers the relation between the two durations given its sensory measurements of duration from each stimulus or stimulus element. We further assume that the biases in sensory measurements are not accessible to the brain at the inference stage. It is very unlikely that the brain learns the true distributions from which the durations are sampled because of the noise in their sensory measurements and the biases introduced by different types of stimuli. For simplicity, we model the belief of the distributions by convolution of the true distributions of the durations (of reference and comparison stimuli) with a Gaussian kernel, as demonstrated in **Figure 3C**. The asymmetric shapes of these distributions result from the logarithmic transformation of duration.

We constructed models by all combinations of assumptions concerning each of three factors: how to form a representation of duration for HL, whether the memory of the sensory measurement of a stimulus' duration decays over time, and how the brain incorporates prior belief of the distributions of duration in their decision. After constructing these models, we performed a thorough factorial model comparison to examine the performance of each hypothesis in each of the three factors (van den Berg et al., 2014).

For the first factor, we considered the "weighting" hypothesis, "optimal integration" hypothesis, "selection" hypothesis, and "reliable stimulus" hypothesis. They differ in how the brain calculates the likelihood of any duration being the true duration, given the sensory measurements of duration based on each elements of HL.



**FIGURE 3 | Model comparison provides quantitative evidence for the “weighting” hypothesis and identified the source of the discrepancy in psychometric curves corresponding to different order of displaying reference and comparison stimuli. (A)** All the conditions tested in Experiment 3. Each condition corresponds to one solid line in the middle, connecting reference, and comparison stimuli. The order in which reference and comparison stimuli were displayed was random. **(B)** The generative model of an example trial for inferring the relation between two durations, if a participant considers the full structure of the task.  $O$ , order of display;  $c-r$ , comparison stimulus was displayed before reference stimulus;  $r-c$ , reference stimulus was displayed before comparison stimulus;  $t_1, t_2$ , durations of the first and second stimuli;  $x_1, x_2$ , sensory measurement of the first and second duration based on the stimuli;  $x_{2,H}, x_{2,L}$ , sensory measurements of the second duration, based on its H and L element, when the stimulus type is HL;  $D$ , decision variable indicating the relation between  $t_1$  and  $t_2$ . **(C)** Illustration of how  $O$  decides the way  $t_1$  and  $t_2$  are sampled from two different distributions corresponding to the reference and comparison stimuli. The colors of the arrows correspond to the respective orders of display  $O$ . **(D)** The workflow of model comparison. Each model is fitted to part of a participant's trials (training data) to find the combination of parameters that maximized the probability of those trials. The fitted parameters are used to predict the behavior in the rest of the participant's trials (testing data). The probability of the testing data assuming the parameters fitted to the training data are logarithmically

transformed to calculate the cross-validated log-likelihood. This procedure is repeated by rotating the selection of testing data over each of the 1/12 portion of the data. Models are compared based on the sum of cross-validated log-likelihood over all the data. **(E)** Average psychometric curves. Figures in the same column correspond to conditions of the same type of reference stimuli. Figures in the same row correspond to the same order of display. Color codes for the type of comparison stimuli. Shaded areas represent the fitted choice probabilities in each condition (mean  $\pm$  s.e.m) by the best model in **(F)**. **(F)** The difference of cross-validated log-likelihood of each model compared to the best model. “weight,” weighting hypothesis; “select,” selection hypothesis; “opt\_int,” optimal integration hypothesis; “reliable\_stim,” reliable stimulus hypothesis; “flat,” flat prior hypothesis; “single,” single prior hypothesis; “double,” double priors hypothesis. **(G)** With individual variability, “weighting” model outperforms each of other models in most participants. The bars represent the differences of the cross-validated log-likelihood of the best models assuming each hypothesis regarding the mechanism of forming the representation of duration for HL stimulus, compared to that of the best model assuming “weighting” hypothesis. A negative bar indicates the model is inferior to the “weighting” model. Each group of bars corresponds to one participant. **(H)** Participants tended to overweight the duration estimate based on H stimulus. The coordinates of each dot correspond to the weight of H estimated in the “weighting” model and the weight of H predicted by the “optimal integration” model for each participant.

For the second factor, we considered two hypotheses. Note that when participants made their judgments on any trial, more time had elapsed since the first stimulus than since the

second stimulus. The first hypothesis, “decay” hypothesis, states that because of the elapse of time, the memory of the first duration decays more than the second, becoming noisier and

more uncertain. To reflect this hypothesis, we assumed that the standard deviation of the sensory measurement of the first duration is scaled up by a constant factor relative to that of the second duration. The second, “no decay” hypothesis, states that the standard deviation is the same regardless of whether a stimulus is presented first or second.

For the third factor, we considered three hypotheses. The first one, the “flat prior” hypothesis, states that the brain does not take into account any prior distribution of duration, thus its judgments are purely based on sensory measurements of duration. The second one, the “single prior” hypothesis, states that the brain learns the mixture of the durations of reference and comparison stimuli as a global distribution and assumes that both durations on any trial are sampled from this distribution. The third one, the “double priors” hypothesis, states that the brain learns the full structure of the generative model in **Figure 3C** that the two durations on any trial are sampled from two different distributions and displayed in random order. Consequently, it incorporates the two learnt distributions and considers both the possible orders of display in the decision process.

The workflow of the model comparison is illustrated in **Figure 3D**. For each model, we derived the decision rules of judging the relation between two durations given any possible sensory measurements on a trial. By integrating the hypothesized distributions of sensory measurements over the range where one of the two judgments should be made according to the decision rule, we obtained the probability that a participant should have made that judgment on any trial (we denote this by choice probability). The choice probability depends on the parameters in each model. Each model thus can be fitted to a subset of data (denoted by training data) of a participant by finding the parameters that maximizes the product of the choice probabilities of all trials in the training data. Each model can be evaluated by predicting the probabilities of the judgments that the participant had made in the rest of the trials (denoted by testing data) based on the parameters fitted to the training data. We conducted 12-fold cross-validation of each model on each participant's data. The logarithm of the product of predicted probabilities over all testing data in the 12-fold cross-validation was compared between models. We denote this measure by cross-validated log-likelihood. This measure is not sensitive to the complexity of the models. A model that is unnecessarily complex would be overfitted to the training data, resulting in low cross-validated log-likelihood.

**Figure 3F** shows the difference of cross-validated log-likelihood of each model from the model that is on average the best across all participants. The more negative the difference is, the worse a model performs. There are several observations from this figure. (1) The largest distinction of model performance was introduced by the assumptions about memory decay and prior belief of duration distribution. Models that assume the existence of memory decay and assume the brain incorporates prior belief of the duration distribution in either form of “single prior” and “double priors” largely outperformed models that do not make these assumptions. By investigating the choice probability predicted by each model, we found that only the combination of the assumptions of memory decay and incorporation of

prior(s) of non-flat form can introduce a difference in choice probability between different orders of displaying reference and comparison stimuli. (2) On average across participants, the “weighting” hypothesis was the best model to describe the representation of duration of the HL stimulus. Among models that can explain the effect of displaying order, the best model was the one assuming a combination of the “weighting” hypothesis, the “decay” hypothesis and the “double priors” hypothesis in the three factors, respectively. Paired *t*-tests between the cross-validated log-likelihood of all other models and that of the best model revealed that the best model outperformed every of other models significantly (The *p*-values passed Holm–Bonferroni multiple comparison thresholds with  $\alpha = 0.05$ . The largest *p*-value was 0.016 when comparing the best model against the model assuming a combination of “optimal integration,” “decay,” and “double priors”). The average difference across participants between the best model and the models with other hypotheses regarding the representation of the duration of HL was at least 3.2 (the best among those models with other hypotheses was the one assuming “selection,” “decay,” and “double priors”). Notice that this difference is in the logarithmic scale. It means that the best model with the “weighting” hypothesis performs at least 25 times as well as models assuming other hypotheses regarding the perceived duration of HL. Since the cross-validated log-likelihood is on the same scale as Bayes factor, the guidance of drawing conclusion on model performance based on Bayes factor (Kass and Raftery, 1995) can help judge the strength of evidence for the best model. According to Kass and Raftery, such difference as observed in the result of Experiment 3 is considered as “strong” evidence for the best model. **Figure 3E** overlays the average psychometric curves over the choice probability fitted by the best model.

**Figure 3G** displays the model performance for each individual participant, focusing on the mechanism of estimating duration of HL. For each participant and for each hypothesis regarding the perceived duration of HL, we identified the best model among the ones with that hypothesis. The difference in cross-validated log-likelihood between each of these best models and the best model with the “weighting” hypothesis is plotted in **Figure 3G** for each participant. Although there is individual difference with respect to the best model for each participant, the “weighting” hypothesis outperforms each of other hypotheses in most participants.

We further compared the estimated weight of H element in the best model with the weight predicted by “optimal integration” based on the standard deviation of the duration estimates of the H and L (**Figure 3H**). The participants' weights of H element ( $0.70 \pm 0.05$ ) were significantly larger than those predicted by “optimal integration” [ $0.50 \pm 0.03$ , paired *t*-test,  $t_{(19)} = 3.53$ ,  $p = 0.002$ ]. There was no significant correlation between weights estimated in the best model and the weights predicted by “optimal integration” ( $p = 0.86$ ).

The discrepancy in psychometric curves found in Experiment 1 can also be accounted for by the same mechanism found in Experiment 3. A model constructed with “decay” and “double-priors” hypotheses fitted well to the psychometric curves (**Figure 4**). Models constructed with “no-decay” or “flat-prior”



hypotheses cannot predict such discrepancy corresponding to different orders of display (figures not shown).

The result of Experiment 3 confirmed that the representation of duration of HL is best described by weighting the duration estimates based on each stimulus element. The brain appears to weight H more than predicted by “optimal integration.” In addition, it shed light on the source of discrepancy in participants’ judgments between different orders of displaying reference and comparison stimuli. Degradation of memory with elapsing time and incorporation of prior distributions of duration jointly account for this discrepancy.

## Discussion

In this study, we first used a two-alternative forced choice task to confirm previous finding that perceived duration is biased by the temporal frequency or speed of a visual stimulus. We further asked how the brain forms a representation of duration when two visual stimuli are displayed simultaneously, one of lower temporal frequency and one of higher temporal frequency. By both qualitatively testing predictions of different models and quantitatively comparing models based on cross-validated log-likelihood, we concluded that the model that best explains the data assumes the duration representation of such joint stimuli is formed by weighting the estimates of duration based on each stimulus element. However, participants’ behavior could not be explained well by the framework of statistically optimal integration. Instead, they tended to overweight the evidence of duration from the stimulus element of higher temporal frequency. In addition, we found that the joint effect of memory decay and incorporation of prior belief of the distributions of duration can account for a discrepancy between psychometric

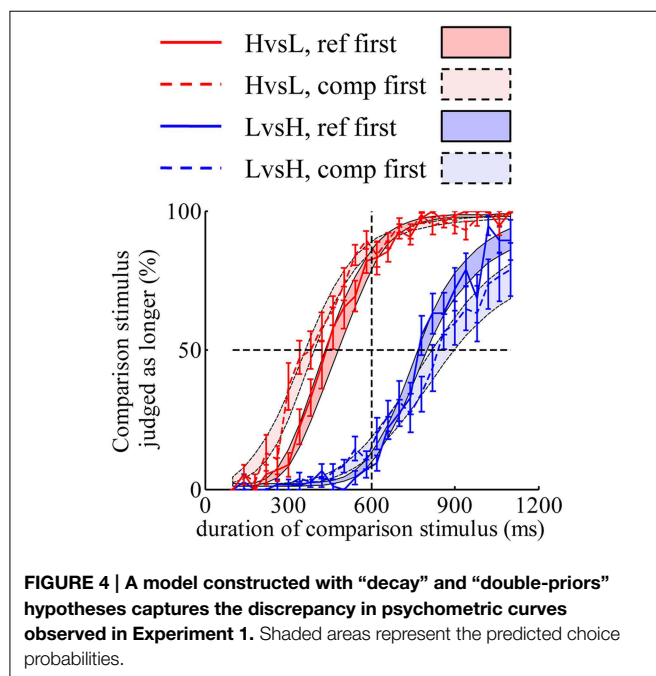
curves of trials belonging to the same condition but with different orders of displaying reference and comparison stimuli.

Previously, the perceived duration of a sequentially concatenated stimulus that is composed of intermittent periods of static and drifting stimuli was found to be perceived shorter than a constantly drifting stimulus of the same duration, but not different from a static stimulus (Bruno et al., 2012). This appears in contrast to our finding that participants overweight the estimate based on the H element when estimating the duration of HL. We should note that in their experiment, the static and drifting intervals of a stimuli were concatenated, rather than presented simultaneously. Therefore, estimating duration of the concatenated stimulus may be viewed as summing the durations of each short interval during which the stimulus was constantly drifting or static instead of averaging the durations of those short intervals. In contrast, the H and L elements in our HL stimulus were displayed simultaneously. Given the large difference in the temporal structures of the stimuli between the two studies, the results of the two studies may not be directly comparable.

In all of our analyses, the curve fitting and modeling were performed after taking logarithmic transformation of duration. This was done because the Weber’s law in duration perception (Gibbon, 1977; Buhusi and Meck, 2005) can be easily captured by assuming a constant level of noise on a logarithmic scale of duration. Fitting a Gaussian cumulative function to the data in Experiment 1 and 2 without logarithmic transformation generated qualitative identical results in all the comparisons critical to our conclusions. We did not attempt to model the data of Experiment 3 on a linear scale of duration because the assumption that sensory measurements follow a Gaussian distribution on a linear scale would result in negative duration estimates, which is meaningless. Additional complexity exists if one chooses to model in linear scale and to assume that the standard deviation of the sensory measurement scales with the duration, because the likelihood function cannot be analytically described by Gaussian function anymore in such a case (Girshick et al., 2011).

In our experiments, we utilized the illusory phenomenon that perceived duration is biased by the temporal frequency or speed of a visual stimulus (Kanai et al., 2006; Kaneko and Murakami, 2009) to manipulate the length of perceived duration without changing the physical duration of a stimulus. There still exists a debate on whether the bias is induced by temporal frequency or speed (Kaneko and Murakami, 2009; Linares and Gorea, 2015). Our result is independent from the answer to this debate, because the spatial frequency was constant in all stimuli and temporal frequency was proportional to speed in our experiments. One may worry that observers could have just used the onsets and offsets to judge duration in our task. This possibility is not compatible to our result because purely judging duration based on the onsets and offsets would not give rise to the difference in perceived duration between H and L, or between HL and the other two types of stimuli.

Several hypotheses have been proposed to account for the influence of temporal frequency or speed on perceived duration. Our results may provide constraints to these hypotheses. First, one hypothesis was that perceived duration may be based



on the amount of change in the environment (Fraisse, 1963; Gibson, 1975; Poynter, 1989; Brown, 1995; Kanai et al., 2006). A quantitative formalization of this idea in the Bayesian observer framework was recently introduced (Ahrens and Sahani, 2011). A second hypothesis was based on the observation that stimuli of longer perceived duration, including those of higher temporal frequencies, typically also elicit larger neural responses. This hypothesis proposed that perceived duration may reflect the neural energy expended to encode sensory stimuli (Pariyadath and Eagleman, 2007; Eagleman and Pariyadath, 2009). Lastly, within the traditional “internal clock” framework of time perception, another hypothesis proposed that fluctuation of neural activity in visual cortex modulated by sensory stimuli may play a role in the tick rate of the clock (Kanai et al., 2006; Kaneko and Murakami, 2009). For the hypothesis based on amount of changes, our results suggest that perceived duration is not based on the total number of changes in all stimuli. Similarly, for the hypothesis based on neural energy, our results suggest that the perceived duration is not formed by summing the neural response to all stimuli, at least for dynamic stimuli. Both of these hypotheses can still be valid if we assume that duration estimates are based on local stimuli and these estimates are further weighted to form a global representation. For the hypothesis within an “internal clock” framework, our results suggest that the clock signals may come from distributed sources in sensory cortex and the tick counts from each source may be fused by weighted average. In contrast, if one assumed there is only one centralized clock, it would be difficult to explain the difference in JNDs when participants compare different types of stimuli. Although our “weighting” hypothesis resembles the spirit of cue integration in the Bayesian observer model, the “optimal integration” hypothesis did not provide the best account for our data.

Note that our implementation of the “optimal integration” hypothesis in Experiment 3 made some simplifying assumptions compared to the modeling framework of Ahrens and Sahani (2011). First, in their paper, the likelihood of duration was calculated as the probability of observing the changes between several samples in a dynamic luminance signal by assuming the signal follows the temporal statistics in natural scenes. By simulating this calculation one can obtain the biases of perceived duration due to different temporal frequencies. We did not use this approach to predict the biases because we found that the bias depends on free parameters such as the number of samples, sampling rates, and the contrast of stimuli compared to that of luminance signals in natural scene. Instead, we simply assumed the biases and standard deviations of the sensory measurements of duration are free parameters for each participant. This simplification should not influence our conclusion as long as the distribution of sensory measurements predicted by simulating their model approximates a Gaussian distribution. Second, in the model of Ahrens and Sahani’s, there was an additional source of duration estimation purely based on internal neural activity, independent from the sensory inputs. We did not include this internal estimation in our models because it was shown that this internal estimation was not crucial to the predictions of their model (Ahrens and Sahani, 2011). However, even if we had

included such an internal estimation, optimal integration should still predict  $\sigma_{HL} \leq \min\{\sigma_H, \sigma_L\}$  in Experiment 2, which was not reflected in the comparison of JNDs.

In Experiment 3, we found that memory decay and incorporation of the prior distributions of duration together account for the discrepancy in the threshold and slope of psychometric curves corresponding to different orders of display. The discrepancy in threshold resembles a phenomenon sometimes called the “time-order error” (Hellstrom, 1985). A similar discrepancy in the slope of psychometric curves was also found in many other studies of perceptual judgments (Nachmias, 2006; Lapid et al., 2008; Bruno et al., 2010, 2012; Ahrens and Sahani, 2011). It was proposed that an implicit standard was used in such comparison (Nachmias, 2006; Lapid et al., 2008). In our minds, this so-called “implicit standard” or “internal standard” plays a similar role as the prior distribution in our “single prior” model. In the model by Lapid et al. (2008), participants only weight the “internal standard” with the sensory evidence of the first stimulus but not with that of the second stimulus. In our models assuming “single prior” and “memory decay,” the decay of memory causes the likelihood function of the first duration to be wider than that of the second. This in turn makes the influence of the prior distribution to the posterior distribution for the first duration stronger than for the second. This is similar to giving more weight to the “internal standard” when calculating a weighted average of the “internal standard” and the sensory estimate of duration. Our modeling result (Figure 3F) suggests that such discrepancy due to the order of display may reflect an optimal strategy to integrate sensory evidence with prior belief of the structure of the task. A similar model was recently proposed to account for an order effect in a task of discriminating lengths of bars (Ashourian and Loewenstein, 2011). The fact that a common mechanism can account for related phenomena in both spatial and timing tasks indicates that similar inference strategies may be used in various domains of perceptual tasks. Here we give an intuitive explanation of why the prior distributions and memory decay jointly causes the effect of the displaying order, taking the “double priors” hypothesis as an example. Under this hypothesis, the brain separately calculates the posterior probabilities of the first duration being longer/shorter than the second based on each hypothetical order of display, and averages these probabilities to make the final judgment. To calculate the posterior probabilities of the relation between the durations, the brain needs to calculate the posterior probabilities of the duration of each stimulus. The prior distribution learnt from the comparison durations is much flatter than that learnt from the standard duration, and is thus less informative. Because it is less informative, it has smaller contribution to the posterior distribution no matter if it is used to infer the duration of a standard stimulus or of a comparison stimulus. On the contrary, the prior distribution corresponding to the standard duration is more concentrated and thus more informative. But it is only beneficial to the accuracy of judgment when it is used to calculate the posterior distribution of the duration for a stimulus that is actually the standard stimulus. If it is used to calculate the posterior distribution of a comparison stimulus, it “drags” the mass of the posterior distribution toward the standard duration, which makes the judgment more difficult.

On the other hand, the relative contribution of the prior distribution to the posterior distribution also depends on the shape of the likelihood function of duration. The prior has relatively stronger impact on the posterior if the likelihood is flatter (less informative). This is the case for a stimulus that appears first in a trial, due to the decay of memory. Therefore, in the trials of which the first stimulus is the standard stimulus, the prior distribution corresponding to the standard duration provides larger benefit for estimating the posterior distribution of the standard duration but generates less “dragging” effect on the posterior distribution of the comparison stimulus. In the trials of which the first stimulus is the comparison stimulus, the “dragging” effect is stronger for the comparison stimulus but the benefit is weaker for the standard stimulus. This explains why the psychometric curve is steeper when the standard stimulus appears first.

One may worry that the order effect may be caused by lower uncertainty of the location of the second stimulus than that of the first. Because the effect of the order of display is observed in many other studies which do not manipulate the location of stimuli as we do, we think the difference in uncertainty of the position of the stimuli is unlikely the major cause of the order effect.

Observers’ behavior in cross-modality cue combination tasks of many spatial features can often be well described by statistically optimal integration or appear close to optimality (Jacobs, 1999; Ernst and Banks, 2002; Battaglia et al., 2003). However, it is puzzling that behavior in cue combination tasks of duration or other temporal features often deviates from optimality in one way or another (Burr et al., 2009; Shi et al., 2010; Hartcher-O’Brien and Alais, 2011; Tomassini et al., 2011). Are brains simply suboptimal when it comes to time? It is difficult to give a comprehensive explanation of the sub-optimality; we can only provide some speculations. The first possibility is the role of causal inference (Knill, 2003, 2007; Körding et al., 2007; Shams and Beierholm, 2010): the brain not only needs to integrate different cues to form a more reliable estimation, but also needs to infer which of the cues may be generated by a different cause and should not be integrated. When two cues conflict too much or their relation violates some constraints, the brain should not integrate them but should instead treat them as from different sources. In spatial cue integration tasks, the temporal contingency between cues provides a strong clue that the cues may be generated from the same source. Unfortunately, in order to study duration cue combination, researchers often have to make the physical durations of the stimuli different (Hartcher-O’Brien and Alais, 2011; Ayhan et al., 2012). This creates asynchrony in onset and offset time between stimuli, which provides a strong clue that they should not be integrated. In fact, Ayhan et al. (2012) found a poorer performance when judging the average duration of multiple asynchronous stimuli than when judging the duration of a single item. They also found no significant difference between judging two items and judging eight items. It is possible that when stimuli are asynchronous, the brain does not perform weighted average but randomly selects one stimulus to estimate duration. Our use of temporal frequency to bias perceived duration avoided this asynchrony. However, it is still possible that the difference between the duration estimates

of the H and L elements may be too large for participants to integrate them on some trials. Future studies that systematically manipulate the temporal frequencies of the two stimuli may help answer whether causal inference is the major cause of the apparent sub-optimality in combining duration estimates. A second possibility is that the stimuli used are not common in the natural environment and the brain may have a wrong belief about the precision of duration estimation based on each type of stimulus. Third, the H element may draw more attention than the L element, and the reliability of duration estimation may be changed due to different levels of attention. Lastly, it is possible that participants may have insufficient knowledge of some task-relevant information. For example, they may have learnt a wrong prior distribution, which may translate to apparent sub-optimality. These possibilities all call for future investigation. We believe that our approach of manipulating perceived duration can be further extended in studying many questions related to the integration of duration estimation.

In our experiments, we only manipulated the bias of perceived duration by temporal frequency, but did not attempt to manipulate the precision of the perceived duration. The difference in the precision of duration estimates of H and L were inherent to each participant. This reflects another limitation in studying cue combination in time perception: to our knowledge, there are few, if any, manipulations of visual stimuli that can independently influence the magnitude and precision of perceived duration (although see Hartcher-O’Brien et al., 2014, where the precision of perceived duration of auditory stimuli was manipulated by the signal to noise ratio of a tone). It is still largely unknown what determines the precision of duration estimation of different types of stimuli, such as the H and L stimuli in our experiments. Understanding how and why variability of duration perception changes with different stimulus features may provide insights into the mechanism by which duration is estimated based on sensory signals. Quantifying the statistics of natural scenes and deriving the optimal encoding and decoding strategy has been a fruitful approach in generating models for how the brain might solve spatial perception tasks. The performances of such models often highly resemble the performance of human observers (Geisler et al., 2009; D’Antona et al., 2013; Burge and Geisler, 2014). Only a few studies in time perception have taken this perspective (Ahrens and Sahani, 2011). We speculate that further analysis of the statistical structure of temporal signals in natural environments may identify the optimal strategy to estimate time based on natural signals and provide ways to understand the variability in duration judgments.

## Data Analysis and Modeling

### Experiment 1

We fitted each participant’s responses by psychometric functions with shapes following Gaussian cumulative distribution. Trials of both orders of display belonging to the same condition were treated equally when fitting a psychometric function to them.

For trials in the LvsH condition, we denote by  $t_{i,L}$  the logarithmic transformation of the physical duration of the comparison stimulus on the  $i$ th trial. Similarly, for trials in the

HvsL condition, we denote by  $t_{i,H}$  the logarithmic transformation of the physical duration of the comparison stimulus on the  $i$ th trial. We assume that the probability of a participant's response  $r_{i,L}$  for the  $i$ th trial of the LvsH condition is

$$p(r_{i,L} = \text{"longer"} \mid t_{i,L}, b_{LvsH}, \sigma_{LvsH}, \lambda) \\ = (1-\lambda) \Phi(t_{i,L} + b_{LvsH}; t_{ref}, \sigma_{LvsH}) + \frac{1}{2}\lambda \quad (9)$$

$$p(r_{i,L} = \text{"shorter"} \mid t_{i,L}, b_{LvsH}, \sigma_{LvsH}, \lambda) \\ = 1 - p(r_{i,L} = \text{"longer"} \mid t_{i,L}, b_{LvsH}, \sigma_{LvsH}, \lambda) \quad (10)$$

Similarly, we assume the probability of response  $r_{i,H}$  for the  $i$ th trial of HvsL condition is

$$p(r_{i,H} = \text{"longer"} \mid t_{i,H}, b_{HvsL}, \sigma_{HvsL}, \lambda) \\ = (1-\lambda) \Phi(t_{i,H} + b_{HvsL}; t_{ref}, \sigma_{HvsL}) + \frac{1}{2}\lambda \quad (11)$$

$$p(r_{i,H} = \text{"shorter"} \mid t_{i,H}, b_{HvsL}, \sigma_{HvsL}, \lambda) \\ = 1 - p(r_{i,H} = \text{"longer"} \mid t_{i,H}, b_{HvsL}, \sigma_{HvsL}, \lambda) \quad (12)$$

where  $\lambda$  is the probability that the participant would make random guess (lapse rate, common for both conditions);  $b_{LvsH}$  is the bias of perceived duration of stimulus L relative to H in the LvsH condition (in the log scale of duration);  $b_{HvsL}$  is the bias of perceived duration of stimulus H relative to L in the HvsL condition;  $\sigma_{LvsH}$  and  $\sigma_{HvsL}$  reflect the sensitivity to duration difference in the two conditions (JND on the logarithmic scale of duration).  $\Phi(\cdot)$  is Gaussian cumulative distribution function.

We assumed the responses are independent between trials.

The likelihood of the parameters  $L(b_{LvsH}, \sigma_{LvsH}, b_{HvsL}, \sigma_{HvsL}, \lambda) = p(\text{data} \mid b_{LvsH}, \sigma_{LvsH}, b_{HvsL}, \sigma_{HvsL}, \lambda)$  could then be calculated by the product of the probability of response for each trial:

$$L(b_{LvsH}, \sigma_{LvsH}, b_{HvsL}, \sigma_{HvsL}, \lambda) \\ = p(\text{data} \mid b_{LvsH}, \sigma_{LvsH}, b_{HvsL}, \sigma_{HvsL}, \lambda) \\ = \prod_{i=1}^N P(r_{i,L} \mid t_{i,L}, b_{LvsH}, \sigma_{LvsH}, \lambda) \cdot \\ \prod_{i=1}^N P(r_{i,H} \mid t_{i,H}, b_{HvsL}, \sigma_{HvsL}, \lambda) \quad (13)$$

where  $N$  is the number of trials in each condition. For each participant, we fitted all the parameters  $b_{LvsH}$ ,  $b_{HvsL}$ ,  $\sigma_{LvsH}$ ,  $\sigma_{HvsL}$ , and  $\lambda$  simultaneously to maximize  $L(b_{LvsH}, \sigma_{LvsH}, b_{HvsL}, \sigma_{HvsL}, \lambda)$ , using the "fmincon" function in Matlab. Since the curve fitting was performed after logarithmic transformation of duration, the bias terms  $b_{LvsH}$  and  $b_{HvsL}$  represent duration distortion in the logarithmic scale. We then calculated  $e^{b_{LvsH}}$  and  $e^{b_{HvsL}}$  as the duration distortion ratio plotted in Figure 1C.

## Experiment 2

The procedure of fitting parameters of psychometric functions was similar to that in Experiment 1. The bias terms  $b_{LvsH}$  and  $b_{HvsL}$  were replaced by  $b_L$ ,  $b_H$ , and  $b_{HL}$ , corresponding to the bias of the perceived duration of each type of comparison stimulus relative to that of the reference stimulus (in the log scale of duration). The JND terms  $\sigma_{LvsH}$  and  $\sigma_{HvsL}$  were replaced by  $\sigma_L$ ,  $\sigma_H$ , and  $\sigma_{HL}$  for each condition.

## Experiment 3 Generative Model

Participants' judgments were considered as an inference process. In Figure 3B, we illustrate an example of the generative models which we assume this inference process may be based on if the brain considers the full structure of the task. On each trial, a binary variable  $O$  determines the order in which the stimuli of different durations are displayed to the participant. With probability of 0.5, the reference stimulus is displayed before the comparison stimulus (we denote this by  $O = \text{"r-c"}$ ). With probability of 0.5, the comparison stimulus is displayed before the reference stimulus (we denote this by  $O = \text{"c-r"}$ ).  $t_1$ , the true duration of the first stimulus, and  $t_2$ , the true duration of the second stimulus, are sampled from the corresponding distributions of reference stimulus and comparison stimulus. Figure 3C illustrates this sampling process. The brain does not have access to the order  $O$  or the true durations  $t_1$  and  $t_2$ . Instead, it has noisy neural measurements of durations that can vary from trial to trial. We denote these measurements by  $x_1$  and  $x_2$ . Here,  $t$  and  $x$  are both in logarithmic scale of duration.

In the cases that the stimulus type in duration  $t_i$  ( $i = 1, 2$ ) is H or L, we assumed that the distribution of  $x_i$  follows a Gaussian distribution on the logarithmic scale of duration. The mean of the distribution is biased by the corresponding stimulus type H or L, as described in Equations (1) and (2).

In the case that the stimulus type in duration  $t_i$  ( $i = 1, 2$ ) is HL, one noisy measurement is generated based on each element of HL. Figure 3B illustrates an example of such a case when the stimulus of duration  $t_2$  is HL. We denote the measurements based on the two elements of HL by  $x_2 = \{x_{2,H}, x_{2,L}\}$ . We assumed that the distribution of duration measurement based on each element is the same as when only that element is displayed, and independent from each other:

$$x_{i,H} \sim N(t + b_H, \sigma_H^2) \quad (i = 1, 2) \quad (14)$$

$$x_{i,L} \sim N(t + b_L, \sigma_L^2) \quad (i = 1, 2) \quad (15)$$

## Inference Process

The brain only has access to  $x_1$  and  $x_2$ . What participants report is their belief of the relation between  $t_1$  and  $t_2$ , denoted by decision variable  $D$  ( $D = 0$  means  $t_1 > t_2$  and  $D = 1$  means  $t_1 < t_2$ ). The process of generating a response about  $D$  based on noisy observations  $x_1$  and  $x_2$  is the inference process that we modeled.

We assumed that the brain estimates the posterior distributions of stimulus durations  $t_1$  and  $t_2$  based on  $x_1$  and  $x_2$ :

$$p(t_i \mid x_i) = \frac{p(x_i \mid t_i) \cdot p(t_i)}{p(x_i)}, \quad (i = 1, 2) \quad (16)$$



The posterior distribution is proportional to two factors:  $p(t_i)$ , the prior distribution of  $t_i$ , and  $p(x_i | t_i)$ , the likelihood of  $t_i$ . The former is a participant's belief of the general distribution of the duration in the experiment without any sensory evidence. The latter is the probability that any particular  $t_i$  can generate the sensory measurement  $x_i$ , regardless of the prior belief.

Based on  $p(t_i | x_i)$ , the brain further calculates the posterior probability of the decision variable  $D$ :

$$p(D = 0 | x_1, x_2) = p(t_1 > t_2 | x_1, x_2) \\ = \int_{-\infty}^{+\infty} dt_1 \int_{-\infty}^{t_1} dt_2 p(t_1 | x_1) p(t_2 | x_2) \quad (17)$$

$$p(D = 1 | x_1, x_2) = p(t_1 < t_2 | x_1, x_2) \\ = \int_{-\infty}^{+\infty} dt_2 \int_{-\infty}^{t_2} dt_1 p(t_1 | x_1) p(t_2 | x_2) \quad (18)$$

If  $p(D = 0 | x_1, x_2) > p(D = 1 | x_1, x_2)$ , the participant reports  $t_1$  as being longer, otherwise he/she reports  $t_2$  as being longer. If Equations (17) and (18) are expanded by plugging in Equation (16), we notice that  $p(x_1)p(x_2)$  is shared in both the formula of  $p(D = 0 | x_1, x_2)$  and  $p(D = 1 | x_1, x_2)$ . Therefore, the terms  $p(x_1)$  and  $p(x_2)$  can be ignored in making judgment about  $D$ .

### Choice Probability

While the inference process described above is deterministic,  $x_1$  and  $x_2$ , the measurements of duration based on certain neural processes in the visual pathway are stochastic. They can vary from trial to trial even if the physical durations are the same. In our modeling, this variation was the major source of variability in participants' judgments. We did not make specific assumption on how  $x_1$  and  $x_2$  are generated. We only made the simple assumption that their distributions follow Equations (1) and (2). In order to calculate the probability that a participant makes a certain judgment on a trial, we integrated over the space of distribution of  $x_1$  and  $x_2$  where the corresponding judgment should be made according to the above decision rule. In addition, similarly as in Experiment 1 and 2, we included a lapse rate term which describes the probability that a participant fails to pay attention to the stimuli and makes a random guess. The choice probability thus takes the following form:

$$p_{M,\theta}(r | t_1, t_2) = \begin{cases} \frac{1}{2}\lambda + (1-\lambda) \int_{-\infty}^{+\infty} dx_1 \int_{-\infty}^{+\infty} dx_2 H(p_{M,\theta}(D = 1 | x_1, x_2) - p_{M,\theta}(D = 0 | x_1, x_2)) \\ \quad \cdot p_{M,\theta}(x_1 | t_1) \cdot p_{M,\theta}(x_2 | t_2), & \text{if } r = \text{"}t_2 \text{ is longer"} \\ \frac{1}{2}\lambda + (1-\lambda) \int_{-\infty}^{+\infty} dx_1 \int_{-\infty}^{+\infty} dx_2 H(p_{M,\theta}(D = 0 | x_1, x_2) - p_{M,\theta}(D = 1 | x_1, x_2)) \\ \quad \cdot p_{M,\theta}(x_1 | t_1) \cdot p_{M,\theta}(x_2 | t_2), & \text{if } r = \text{"}t_2 \text{ is shorter"} \end{cases} \quad (19)$$

In the above equation,  $r$  is the judgment.  $M$  indicates the model under consideration.  $\theta$  represents all the free parameters of model  $M$ .  $H(\cdot)$  means a step function which outputs 1 only when the input is larger or equal to 0 and outputs 0 otherwise.  $\lambda$  is the lapse rate.

An analytic form of the choice probability does not exist as function of  $t_1$  and  $t_2$ . To calculate the integral numerically, we used a Gaussian-Hermite quadrature of order 7 to approximate

the integration over  $x_1$ . For a value of  $x_1$  chosen as the abscissa in the integration, the value of  $x_2$  that satisfies  $p(D = 0 | x_1, x_2) = p(D = 1 | x_1, x_2)$  was found by numerical search. The step function  $H(\cdot)$  is 1 on one side of this value of  $x_2$  and 0 on the other side. Therefore, the integration over  $x_2$  was calculated based on the cumulative distribution function of  $p(x_2 | t_2)$  at this value of  $x_2$ .

### Model Comparison

Our goal was to understand how the brain forms a duration representation when multiple stimuli, each providing conflicting evidence of duration occur simultaneously. In our modeling framework, the process of forming duration representation based on multiple stimuli is the process of calculating the likelihood of a duration  $t$  when the stimulus is HL. Thus, one major difference between the models under consideration is in their likelihood function  $p(x_{i,L}, x_{i,H} | t_i)$  ( $i = 1, 2$ ), when the stimulus in  $t_i$  is HL and separate sensory measurements  $x_{i,L}$  and  $x_{i,H}$  are formed. In addition, we also aimed to understand the discrepancy observed in the psychometric curves corresponding to different orders of displaying the reference and comparison stimuli. We considered two possible causes for the discrepancy: the sensory measurement of the first duration on a trial may be degraded more than that of the second due to decay of memory, and participants may incorporate the prior belief of duration distribution into their inference process.

Therefore, we constructed models based on three factors: the likelihood function of duration when the stimulus is HL, whether memory decay exists, and how participants incorporate prior belief of stimulus duration during inference.

### Likelihood function

The form of the likelihood function of duration  $t$  when the stimulus is H or L is shared among all models. As the distribution of measurement  $x$  has a constant level of noise over the range of  $t$  (on log scale), a reasonable assumption is that the likelihood function follows the shape of Gaussian function with the same standard deviation as the level of noise:

$$L(t_i) = p(x_i | t_i) = \begin{cases} N(x_i, \sigma_H), & \text{if H stimulus is displayed} \\ N(x_i, \sigma_L), & \text{if L stimulus is displayed} \end{cases} \quad (20)$$

In the above equation, we also assumed that the biases  $b_H$  and  $b_L$  in the distributions of  $x_H$  or  $x_L$ , as in Equation (1) and (2), are not accessible by the brain at the inferring stage. This assumption and the difference between  $b_H$  and  $b_L$  explain why H is judged as longer than L in our modeling framework.

The likelihood function of duration  $t$  when the stimulus is HL differs between models.

In models assuming the “weighting” hypothesis, we assume that the brain first weights the two sensory measurements of duration by Equation (3). The likelihood function of  $t$  is then calculated based on  $x_{HL}$ :

$$p(x_{i,L}x_{i,H} | t_i) = L_{\text{weighting}}(t_i) \\ = N(t_i; w_H x_{i,H} + (1 - w_H)x_{i,L}, \sqrt{w_H^2 \sigma_{i,H}^2 + (1 - w_H)^2 \sigma_{i,L}^2}) \quad (21)$$

We modeled the standard deviation of the likelihood function as in the above equation because it matches the standard deviation of the distribution of  $x_{HL}$  following the weighting scheme in Equation (3).

In models assuming the “optimal integration” hypothesis, a stronger version of the “weighting” hypothesis, the likelihood is the product of the likelihood of  $t$  based on each individual stimulus element, which amounts to:

$$p(x_{i,L}x_{i,H} | t_i) = L_{\text{optimal}}(t_i) = N(x_{i,H}, \sigma_H) \cdot N(x_{i,L}, \sigma_L)$$

In models assuming the “selection” hypothesis, the likelihood function is based only on the stimulus element that is selected to estimate duration:

$$p(x_{i,L}x_{i,H} | t_i) = L_{\text{selection}}(t_i) \\ = \begin{cases} N(x_{i,H}, \sigma_H), & \text{if stimulus H is selected} \\ N(x_{i,L}, \sigma_L), & \text{if stimulus L is selected} \end{cases} \quad (22)$$

In models assuming the “reliable stimulus” hypothesis, the likelihood function is based on the stimulus element which the participants has a smaller standard deviation in his/her estimation of duration:

$$p(x_{i,L}x_{i,H} | t_i) = L_{\text{reliable stimulus}}(t_i) \\ = \begin{cases} N(x_{i,H}, \sigma_H), & \text{if } \sigma_H < \sigma_L \\ N(x_{i,L}, \sigma_L), & \text{if } \sigma_H > \sigma_L \end{cases} \quad (23)$$

In models assuming the “weighting,” “optimal integration,” or “reliable stimulus” hypothesis, the likelihood function can be plugged into the inference process and the choice probability can be calculated for each combination of model parameters.

In models assuming the “selection” hypothesis, if the reference stimulus is HL and the comparison stimulus is H or L, then the two choice probabilities, corresponding to either H or L element being selected from the reference stimulus, are first calculated by plugging the likelihood function corresponding to that stimulus being selected into the inference process. Then these probabilities are further multiplied by the probabilities of H or L being selected and summed together, to calculate the expected choice probability for a given trial.

$$p(r | t_1, t_2, \theta, M) = p_{\text{select H}}(r | t_1, t_2, \theta, M) c_H \\ + p_{\text{select L}}(r | t_1, t_2, \theta, M) (1 - c_H) \quad (24)$$

If the comparison stimulus is also HL, then the equation above is used to first calculate the choice probabilities of either H or L element being selected from the comparison stimulus. They are further multiplied by  $c_H$  and  $1 - c_H$  and summed similarly.

## Memory decay

In order to make a comparison of duration, participants need to hold the memory of the duration of the first stimulus until making judgment. At the time of making judgment, more time had elapsed since the first stimulus than since the second stimulus. It is possible that the representation of duration of the first stimulus was more variable than that of the second stimulus due to decay of memory. Therefore, the second factor that we consider in constructing models is whether the standard deviation of  $x_1$  increases compared to  $x_2$  due to memory decay.

In models assuming the “decay” hypothesis, the standard deviation of the distribution of  $x_1$  is scaled up by a constant  $m$  ( $m > 1$ ) relative to that of  $x_2$  of the same type of stimulus.  $m$  is a free parameter common to all stimulus types. The standard deviation of the likelihood function of the first duration  $t_1$  is also scaled up by  $m$ .

In models assuming the “no decay” hypothesis, there is no difference in the standard deviation of the distributions of  $x_1$  and  $x_2$ , which is equivalent to fixing  $m$  as 1.

## Incorporation of prior distribution

The distribution of duration presented in the experiment was not uniform. It is possible that the brain can gradually learn the distribution of duration as the experiment continues. Furthermore, as illustrated in **Figures 3B,C**, the physical durations of the two stimuli in any trial were sampled from two different distributions with random orders. The brain might further learn this structure. Therefore, we considered three different hypotheses of how the brain might form a belief of the prior distribution of duration.

In models assuming the “flat prior” hypothesis, the brain does not learn any distribution from the experiment but instead assumes any duration is equally possible to occur for both the first and second stimuli. This is equivalent to saying that the posterior of duration is the same as the likelihood of duration:  $p(t_i | x_i) = p(x_i | t_i)$ . The generative model assumed by the brain would be without the parameter of displaying order  $O$  in **Figure 3B**.

In models assuming the “single prior” hypothesis, the brain forms a belief that all stimulus durations are sampled from the same distribution, which is the mixture of the distribution of the reference and comparison duration. Note that it is impossible for participants to learn the exact distribution of the physical duration, because of the noise in sensory measurement of duration, and because H and L type of stimuli cast different biases on the measurements. Therefore, the prior distribution learnt by the brain should be a smoothed version of the true distribution. For simplicity, we assume that the prior distribution  $p(t_i)$  in Equation (16) takes the form of the convolution of a Gaussian kernel with the mixture of distributions of the true duration of both the reference and comparison stimuli.

In models assuming the “double priors” hypothesis, the brain learns the correct generative model as in **Figure 3C**, that durations are sampled from two distributions and a top-level variable  $O$  determines the order in which the two durations are drawn from these distributions. In order to account for both the possible orders of display, the brain separately calculates the posterior probabilities of the decision variable  $D$  based on each

possible order  $O$ , and marginalize over  $O$  by taking the average of these two probabilities:

$$p(D = 0 \mid x_1, x_2) = \frac{p(t_1 > t_2 \mid x_1, x_2, O = \text{"c-r"}) + p(t_1 > t_2 \mid x_1, x_2, O = \text{"r-c"})}{2} \quad (25)$$

$$p(D = 1 \mid x_1, x_2) = \frac{p(t_1 < t_2 \mid x_1, x_2, O = \text{"c-r"}) + p(t_1 < t_2 \mid x_1, x_2, O = \text{"r-c"})}{2} \quad (26)$$

In the above equations,  $p(t_1 > t_2 \mid x_1, x_2, O)$  and  $p(t_1 < t_2 \mid x_1, x_2, O)$  were calculated similarly as in Equation (17), except that the posterior probabilities of  $t_1$  and  $t_2$  depend on the variable  $O$ . We named the prior probability of the duration of the comparison stimuli by  $p_c(t)$ , and that of the reference stimuli by  $p_r(t)$ . The posterior probabilities of  $t_1$  and  $t_2$  corresponding to the two orders of display are:

$$p(t_1 \mid x_1, x_2, O = \text{"c-r"}) = \frac{p_c(t_1)p(x_1 \mid t_1)}{p(x_1)},$$

$$p(t_2 \mid x_1, x_2, O = \text{"c-r"}) = \frac{p_r(t_2)p(x_2 \mid t_2)}{p(x_2)} \quad (27)$$

$$p(t_1 \mid x_1, x_2, O = \text{"r-c"}) = \frac{p_r(t_1)p(x_1 \mid t_1)}{p(x_1)},$$

$$p(t_2 \mid x_1, x_2, O = \text{"r-c"}) = \frac{p_c(t_2)p(x_2 \mid t_2)}{p(x_2)} \quad (28)$$

## References

- Aaen-Stockdale, C., Hotchkiss, J., Heron, J., and Whitaker, D. (2011). Perceived time is spatial frequency dependent. *Vision Res.* 51, 1232–1238. doi: 10.1016/j.visres.2011.03.019
- Ahrens, M. B., and Sahani, M. (2011). Observers exploit stochastic models of sensory change to help judge the passage of time. *Curr. Biol.* 21, 200–206. doi: 10.1016/j.cub.2010.12.043
- Ashourian, P., and Loewenstein, Y. (2011). Bayesian inference underlies the contraction bias in delayed comparison tasks. *PLoS ONE* 6:e19551. doi: 10.1371/journal.pone.0019551
- Ayhan, I., Bruno, A., Nishida, S., and Johnston, A. (2009). The spatial tuning of adaptation-based time compression. *J. Vis.* 9, 2.1–12. doi: 10.1167/9.11.2
- Ayhan, I., Bruno, A., Nishida, S., and Johnston, A. (2011). Effect of the luminance signal on adaptation-based time compression. *J. Vis.* 11:22. doi: 10.1167/11.7.22
- Ayhan, I., Revina, Y., Bruno, A., and Johnston, A. (2012). Duration judgments over multiple elements. *Front. Psychol.* 3:459. doi: 10.3389/fpsyg.2012.00459
- Battaglia, P. W., Jacobs, R. A., and Aslin, R. N. (2003). Bayesian integration of visual and auditory signals for spatial localization. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* 20, 1391–1397. doi: 10.1364/JOSAA.20.001391
- Birngruber, T., Schröter, H., and Ulrich, R. (2014). Duration perception of visual and auditory oddball stimuli: does judgment task modulate the temporal oddball effect? *Atten. Percept. Psychophys.* 76, 814–828. doi: 10.3758/s13414-013-0602-2
- Brainard, D. H. (1997). The psychophysics toolbox. *Spat. Vis.* 10, 433–436.
- Brown, S. W. (1995). Time, change, and motion: the effects of stimulus movement on temporal perception. *Percept. Psychophys.* 57, 105–116.
- Bruno, A., Ayhan, I., and Johnston, A. (2010). Retinotopic adaptation-based visual duration compression. *J. Vis.* 10:30. doi: 10.1167/10.10.30
- Bruno, A., Ayhan, I., and Johnston, A. (2012). Effects of temporal features and order on the apparent duration of a visual stimulus. *Front. Psychol.* 3:90. doi: 10.3389/fpsyg.2012.00090
- Buhusi, C. V., and Meck, W. H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nat. Rev. Neurosci.* 6, 755–765. doi: 10.1038/nrn1764
- Burge, J., and Geisler, W. S. (2014). Optimal disparity estimation in natural stereo images. *J. Vis.* 14:1. doi: 10.1167/14.2.1
- Burr, D., Banks, M. S., and Morrone, M. C. (2009). Auditory dominance over vision in the perception of interval duration. *Exp. Brain Res.* 198, 49–57. doi: 10.1007/s00221-009-1933-z
- Carrozzo, M., and Lacquaniti, F. (2012). Effects of speeding up or slowing down animate or inanimate motions on timing. *Exp. Brain Res.* 224, 581–590. doi: 10.1007/s00221-012-3338-7
- D'Antona, A. D., Perry, J. S., and Geisler, W. S. (2013). Humans make efficient use of natural image statistics when performing spatial interpolation. *J. Vis.* 13:11. doi: 10.1167/13.14.11
- Eagleman, D. M. (2008). Human time perception and its illusions. *Curr. Opin. Neurobiol.* 18, 131–136. doi: 10.1016/j.conb.2008.06.002
- Eagleman, D. M., and Pariyadath, V. (2009). Is subjective duration a signature of coding efficiency? *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 1841–1851. doi: 10.1098/rstb.2009.0026
- Ernst, M. O., and Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415, 429–433. doi: 10.1038/415429a
- Fraisse, P. (1963). *The Psychology of Time*. New York, NY: Harper & Row.
- Geisler, W. S., Najemnik, J., and Ing, A. D. (2009). Optimal stimulus encoders for natural tasks. *J. Vis.* 9, 17.1–16. doi: 10.1167/9.13.17
- Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychol. Rev.* 84, 279–325.

As described above, we considered three factors: the mechanism of combining duration estimates based on simultaneous stimuli, the existence of memory decay, and the form of prior distribution. Each combination of these three factors generates one model. We compared 24 models ( $4 \times 2 \times 3$ ) in total based on cross-validated log-likelihoods of the models (van den Berg et al., 2014). We first separated the trials of each participant into 12 subsets. Each subsets contained approximately an equal number of trials belonging to each condition and each order of display (we say “approximately” because the total number of trials is not a multiple of 12). Then for each model, we performed 12-fold cross validation. In each case, we left one subset of trials out as testing data. Trials of the other 11 subsets were treated as training data. We fitted the model to the training data by searching for a combination of parameters that maximizes the product of the choice probabilities over all trials in the training data. Then with parameters fitted to the training data, we calculated the log-likelihood of the testing data as the logarithm of the product of the choice probabilities over all trials in the testing data. The sum of the log-likelihoods of the testing data over the 12 instances of cross-validation is the cross-validated log-likelihood of the model being compared. **Figure 3D** illustrate this procedure.

## Acknowledgments

We would like to thank Dr. Wei Ji Ma for the discussion of Bayesian modeling and Dr. Yael Niv for the suggestion on model comparison.

- Gibson, J. J. (1975). "Events are perceivable but time is not," in *The study of Time II*, eds J. T. Fraser and N. Lawrence (Berlin; Heidelberg; New York, NY: Springer), 295–301.
- Girshick, A. R., Landy, M. S., and Simoncelli, E. P. (2011). Cardinal rules: visual orientation perception reflects knowledge of environmental statistics. *Nat. Neurosci.* 14, 926–32. doi: 10.1038/nn.2831
- Grondin, S. (2010). Timing and time perception: a review of recent behavioral and neuroscience findings and theoretical directions. *Atten. Percept. Psychophys.* 72, 561–82. doi: 10.1038/nn.2831
- Hartcher-O'Brien, J., and Alais, D. (2011). Temporal ventriloquism in a purely temporal context. *J. Exp. Psychol. Hum. Percept. Perform.* 37, 1383–1395. doi: 10.1037/a0024234
- Hartcher-O'Brien, J., Di Luca, M., and Ernst, M. O. (2014). The duration of uncertain times: audiovisual information about intervals is integrated in a statistically optimal fashion. *PLoS ONE* 9:e89339. doi: 10.1371/journal.pone.0089339
- Hellstrom, A. (1985). The time-order error and its relatives: mirrors of cognitive processes in comparing. *Psychol. Bull.* 97, 35–61.
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* 6, 65–70.
- Ivry, R. B., and Schlerf, J. E. (2008). Dedicated and intrinsic models of time perception. *Trends Cogn. Sci.* 12, 273–280. doi: 10.1016/j.tics.2008.04.002
- Jacobs, R. A. (1999). Optimal integration of texture and motion cues to depth. *Vision Res.* 39, 3621–3629.
- Johnston, A., Arnold, D. H., and Nishida, S. (2006). Spatially localized distortions of event time. *Curr. Biol.* 16, 472–479. doi: 10.1016/j.cub.2006.01.032
- Kanai, R., Paffen, C. L., Hogendoorn, H., and Verstraten, F. A. (2006). Time dilation in dynamic visual display. *J. Vis.* 6, 1421–1430. doi: 10.1167/6.12.8
- Kaneko, S., and Murakami, I. (2009). Perceived duration of visual motion increases with speed. *J. Vis.* 9:14. doi: 10.1167/9.7.14
- Kass, R. E., and Raftery, A. E. (1995). Bayes factors. *J. Am. Stat. Assoc.* 90, 773–795. doi: 10.1080/01621459.1995.10476572
- Kleiner, M., Brainard, D., and Pelli, D. (2007). What's new in Psychtoolbox-3? *Perception* 36. ECVF Abstract Supplement.
- Kline, S. R., and Reed, C. L. (2013). Contextual influences of dimension, speed, and direction of motion on subjective time perception. *Atten. Percept. Psychophys.* 75, 161–167. doi: 10.3758/s13414-012-0370-4
- Knill, D. C. (2003). Mixture models and the probabilistic structure of depth cues. *Vision Res.* 43, 831–854. doi: 10.1016/S0042-6989(03)00003-8
- Knill, D. C. (2007). Robust cue integration: a Bayesian model and evidence from cue-conflict studies with stereoscopic and figure cues to slant. *J. Vis.* 7, 5.1–24. doi: 10.1167/7.7.5
- Knill, D. C., and Pouget, A. (2004). The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci.* 27, 712–719. doi: 10.1016/j.tins.2004.10.007
- Körding, K. P., Beierholm, U., Ma, W. J., and Quartz, S. (2007). Causal inference in multisensory perception. *PLoS ONE* 2:e943. doi: 10.1371/journal.pone.0000943
- Lapid, E., Ulrich, R., and Rammsayer, T. (2008). On estimating the difference limen in duration discrimination tasks: a comparison of the 2AFC and the reminder task. *Percept. Psychophys.* 70, 291–305. doi: 10.3758/PP.70.2.291
- Linares, D., and Gorea, A. (2015). Temporal frequency of events rather than speed dilates perceived duration of moving objects. *Sci. Rep.* 5:8825. doi: 10.1038/srep08825
- Long, G. M., and Beaton, R. J. (1980). The contribution of visual persistence to the perceived duration of brief targets. *Percept. Psychophys.* 28, 422–430. doi: 10.3758/BF03204886
- Long, G. M., and Beaton, R. J. (1981). The effects of stimulus numerosity, retinal location, and rod contrast on perceived duration of brief visual stimuli. *Percept. Psychophys.* 29, 389–394. doi: 10.3758/BF03207349
- Merchant, H., Harrington, D. L., and Meck, W. H. (2013). Neural basis of the perception and estimation of time. *Annu. Rev. Neurosci.* 36, 313–336. doi: 10.1146/annurev-neuro-062012-170349
- Nachmias, J. (2006). The role of virtual standards in visual discrimination. *Vision Res.* 46, 2456–2464. doi: 10.1016/j.visres.2006.01.029
- Pariyadath, V., and Eagleman, D. (2007). The effect of predictability on subjective duration. *PLoS ONE* 2:e1264. doi: 10.1371/journal.pone.0001264
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat. Vis.* 10, 437–442.
- Poynter, D. (1989). Judging the duration of time intervals: a process of remembering segments of experience. *Adv. Psychol.* 59, 305–331.
- Schindel, R., Rowlands, J., and Arnold, D. H. (2011). The oddball effect: perceived duration and predictive coding. *J. Vis.* 11:17. doi: 10.1167/11.2.17
- Shams, L., and Beierholm, U. R. (2010). Causal inference in perception. *Trends Cogn. Sci.* 14, 425–432. doi: 10.1016/j.tics.2010.07.001
- Shi, Z., Chen, L., and Müller, H. J. (2010). Auditory temporal modulation of the visual Ternus effect: the influence of time interval. *Exp. Brain Res.* 203, 723–735. doi: 10.1007/s00221-010-2286-3
- Singh, K. D., Smith, A. T., and Greenlee, M. W. (2000). Spatiotemporal frequency and direction sensitivities of human visual areas measured using fMRI. *Neuroimage* 12, 550–564. doi: 10.1006/nimg.2000.0642
- Tomassini, A., Gori, M., Burr, D., Sandini, G., and Morrone, M. C. (2011). Perceived duration of visual and tactile stimuli depends on perceived speed. *Front. Integr. Neurosci.* 5:51. doi: 10.3389/fnint.2011.00051
- Treisman, M. (1963). Temporal discrimination and the indifference interval. Implications for a model of the "internal clock." *Psychol. Monogr.* 77, 1–31.
- Tse, P. U., Intriligator, J., Rivest, J., and Cavanagh, P. (2004). Attention and the subjective expansion of time. *Percept. Psychophys.* 66, 1171–1189. doi: 10.3758/BF03196844
- van den Berg, R., Awh, E., and Ma, W. J. (2014). Factorial comparison of working memory models. *Psychol. Rev.* 121, 124–149. doi: 10.1037/a0035234
- van Wassenhove, V., Buonomano, D. V., Shimojo, S., and Shams, L. (2008). Distortions of subjective time perception within and across senses. *PLoS ONE* 3:e1437. doi: 10.1371/journal.pone.0001437
- Wichmann, F. A., and Hill, N. J. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. *Percept. Psychophys.* 63, 1293–1313. doi: 10.3758/BF03194544
- Xuan, B., Zhang, D., He, S., and Chen, X. (2007). Larger stimuli are judged to last longer. *J. Vis.* 7, 2.1–5. doi: 10.1167/7.10.2

**Conflict of Interest Statement:** 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.

Copyright © 2015 Cai and Eagleman. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# The influence of stimulus repetition on duration judgments with simple stimuli

Teresa Birngruber\*, Hannes Schröter and Rolf Ulrich

Department of Psychology, University of Tübingen, Tübingen, Germany

## OPEN ACCESS

### Edited by:

Lihan Chen,  
Peking University, China

### Reviewed by:

William Matthews,  
University of Cambridge, UK  
Tsuyoshi Kuroda,  
Kyushu University, Japan

### \*Correspondence:

Teresa Birngruber,  
Cognition and Perception,  
Department of Psychology, University  
of Tübingen, Schleichstr. 4,  
72076 Tübingen, Germany  
teresa.birngruber@uni-tuebingen.de

### Specialty section:

This article was submitted to  
Perception Science,  
a section of the journal  
Frontiers in Psychology

**Received:** 26 June 2015

**Accepted:** 30 July 2015

**Published:** 18 August 2015

### Citation:

Birngruber T, Schröter H and Ulrich R  
(2015) The influence of stimulus  
repetition on duration judgments with  
simple stimuli. *Front. Psychol.* 6:1213.  
doi: 10.3389/fpsyg.2015.01213

Two experiments investigated the effects of stimulus repetition vs. stimulus novelty on perceived duration. In a reminder task, a standard and a comparison stimulus were presented consecutively in each trial, and the comparison was either a repetition of the standard or a different stimulus. Pseudowords (Experiment 1) or strings of consonants (Experiment 2) were used as stimuli and the inter-stimulus interval (ISI) between the standard and the comparison was either constant or variable. Participants were asked to judge whether the comparison was shorter or longer than the standard. In both experiments, we observed shorter judged durations for repeated than for novel comparisons whereas the manipulation of the ISI had no pronounced effects on duration judgments. The finding of shorter duration judgments for repeated as compared to novel nonwords replicates the results of a previous study (Matthews, 2011) which employed highly complex stimulus material. The present study shows that changes of simple, semantically meaningless stimuli are sufficient to result in a shorter perceived duration of repeated as compared to novel stimuli.

**Keywords:** time perception, repetition, novelty, nonwords, duration judgment

## 1. Introduction

Human time perception is known to be influenced by many non-temporal aspects (Eagleman, 2008; Grondin, 2010). For example, perceived duration not only depends on physical time but also on the sensory modality stimuli are presented in Goldstone and Lhamon (1974), Wearden et al. (1998), low-level stimulus features (such as contrast: Matthews et al., 2011, or stimulus size: Thomas and Cantor, 1976; Rammsayer and Verner, 2015), and the emotional context of stimulus presentation (Droit-Volet et al., 2011).

Another context effect is the so-called temporal oddball effect. It describes the phenomenon that the duration of deviant stimuli (oddballs) within a stream of homogenous standards is commonly judged as being longer than the duration of the standards (Tse et al., 2004; Pariyadath and Eagleman, 2007; Chen and Yeh, 2009; New and Scholl, 2009; Schindel et al., 2011; Kim and McAuley, 2013; Birngruber et al., 2014). This result from the “stream-based” oddball paradigm has mostly been interpreted as a temporal overestimation of oddballs. But since only relative judgments between standards and oddballs are required, it could just as well reflect a temporal underestimation of standards (see Birngruber et al., 2015, for a study including judgments of standards as well as oddballs).

Matthews (2011) has shown that even a single repetition of a stimulus can result in a shortened judged duration of this stimulus as compared to a novel stimulus. In his experiments, only two

stimuli, first a standard and then a comparison, were presented in each trial and had to be compared in duration (reminder task, see also Ulrich et al., 2006). Naturalistic photographs of different content, e.g., social scenes, nature, objects, and buildings were used as stimuli. The comparison could either be a repetition of the standard or a novel photograph (never encountered before). Matthews observed that repeated comparisons were systematically underestimated compared to novel ones.

The results by Matthews (2011) provide evidence that a single stimulus repetition influences duration judgments. The stimulus material used in this study was rather complex and thus differed on many levels of information: content, categories, color, texture, contrast, etc. Consequently, all these features remained the same between standards and repeated comparisons whereas there were multi-level differences between standards and novel comparisons. In order to examine whether this effect persists even without high-level information, we designed a conceptual replication of Matthews' study using nonwords as stimuli. Nonwords are much simpler than photographs and the only difference for repeated as compared to novel comparisons is whether the letter string of the nonword is repeated or changed. Whether the letter string itself represents rather a low- or a high-level feature is not easy to decide. On the one hand, individual letters obviously vary in shape and nonwords might therefore differ slightly in spatial frequency and overall luminance. On the other hand, many low-level features of nonwords can be easily controlled (e.g., size, color, contrast) and nonwords have per definition no semantic meaning. We chose nonwords as stimuli because high-level information and low-level differences could be minimized while a straight-forward manipulation of repetition was possible. If repetition as compared to a change of information is sufficient to influence perceived duration even if semantic meaning is absent and low-level differences are minimized, repeated nonwords should be judged as being shorter than novel ones.

It should be noted, however, that Matthews recently replicated the repetition effect for a more abstract set of stimuli himself (Matthews, 2015). In Experiments 5 and 6 of this study, nine icons of abstract two-color patterns were combined to  $3 \times 3$  grids and presented as standards and comparisons in a reminder task. While these stimuli had no semantic meaning either, they still contained color, luminance, and shape changes and therefore might have differed on multiple levels. In contrast, the present study examined whether the repetition effect would even generalize to nonword stimuli which are composed of over-learned elements (i.e., letters) and differ only minimally in low-level features.

Furthermore, to address a different issue, we manipulated whether the inter-stimulus intervals between the presentation of the standard and the comparison were predictable or not. Tse et al. (2004) have argued that a fixed temporal structure within a trial might induce rhythm which could interact with duration perception. To investigate this possibility, we presented constant inter-stimulus intervals (as in Matthews, 2011, 2015) in one half of the experiment and variable inter-stimulus intervals (as in Tse et al., 2004) in the other half. If a strictly predictable temporal structure would facilitate rhythmic

processing, temporal discrimination sensitivity should be better with constant than with variable inter-stimulus intervals.

## 2. Experiment 1

### 2.1. Method

#### 2.1.1. Participants

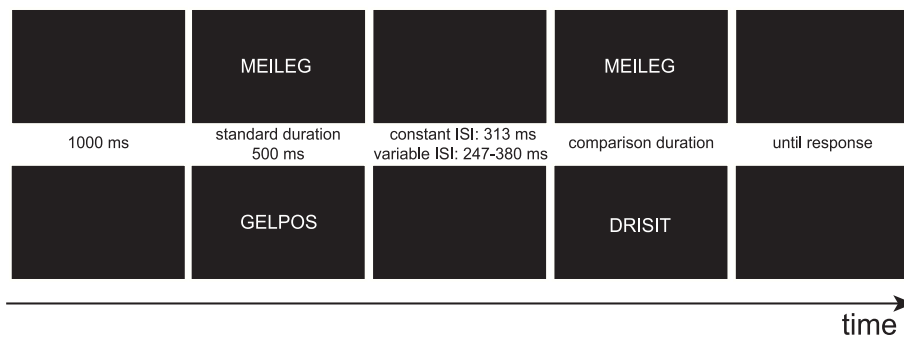
The data of 32 volunteers (22 female, 29 right-handed), aged between 21 and 51 years ( $M = 24.2$  years) entered the analyses. All participants had normal or corrected-to-normal vision, and all received course credit. The experimental session lasted approximately 40 min. Eight additional participants took part in the experiment but had to be excluded due to DL (difference limen) measures larger than 200 ms in at least one of the four conditions. Since the corresponding psychometric functions were almost flat, the PSE estimates were rather unreliable and we therefore considered these data sets to be uninformative with respect to the research question. All participants gave informed consent.

#### 2.1.2. Apparatus and Stimuli

The experiment was programmed in MATLAB® using the Psychophysics Toolbox extension (Brainard, 1997; Pelli, 1997) and presented via a PC with standard VGA monitor ( $1024 \times 768$  pixels, 150 Hz). As nonword stimuli, 104 pseudowords (pronounceable but meaningless letter strings) were taken from the *Verbaler Lerntest* (verbal learning test, Sturm and Willmes, 1999). These pseudowords were the low-associative subset of the items used in this memory test (see Sturm and Willmes, 1999). This means that they were rated as being unlikely to be associated with actual German words. The pseudowords were comprised of six letters and two syllables, e.g., "MEILEG," "DRISIT," or "GELPOS." All pseudowords were presented in capital letters in white font color on a black background, were about 1.8 cm long ( $2.6^\circ$  of visual angle), and were always presented in the center of the screen. For each participant, 81 items were randomly selected from the pool of 104 pseudowords. The "X" and "M" keys of a standard German keyboard served as response keys.

#### 2.1.3. Procedure

The experiment was run in a sound-attenuated, dimly illuminated room. An illustration of the trial structure can be found in **Figure 1**. Each trial started with a blank black screen which was presented for 1000 ms. Then, two stimuli were presented one at a time. The first stimulus (standard) was always presented for 500 ms, the second stimulus (comparison) was presented for one of nine comparison durations: 313, 360, 407, 453, 500, 547, 593, 640, or 687 ms. The two stimuli were separated by an inter-stimulus interval (ISI). In the constant ISI condition, the ISI was 313 ms; in the variable ISI condition, ISIs were randomly selected from the following five durations: 247, 280, 313, 346, and 380 ms. In half of the trials, the comparison was identical to the standard (repeated condition), whereas in the other half of the trials, a different pseudoword was shown (novel condition). The participants were instructed to make a judgment about whether the comparison was shorter or longer than the



**FIGURE 1 | Trial structure of Experiment 1.** An example of a repeated (novel) trial can be seen at the top (bottom). The inter-stimulus interval was either fixed in the constant ISI condition or varied in the variable ISI condition.

Comparison durations varied across nine durations between 313 and 687 ms and the participants' task was to judge whether the comparison was shorter or longer than the standard.

standard, irrespective of condition. After the participant's key press, the next trial started.

The experiment was segmented in one practice block and 12 experimental blocks. The practice block was comprised of 12 trials (six repeated and six novel trials). The items for the practice block were chosen randomly from the pool of pseudowords. Six of the 12 experimental blocks realized the constant ISI condition while the other six blocks realized the variable ISI condition. The ISI condition was blocked and the practice block was of the same ISI condition as the first half of the experiment. Each experimental block was comprised of 54 trials (27 repeated and 27 novel trials), thus each block included all 81 items that were pre-selected for each participant. The same 81 items were presented in each block, but whether individual items appeared in a repeated or novel trial was randomized. The fact that items were therefore presented several times throughout the experiment should not be problematic as repetition effects on time perception seem to be quite short-lived (see Matthews, 2011, Experiment 2 and Matthews, 2015, Experiments 5 and 6). Short breaks were included between experimental blocks and once within each block (every 27 trials). In total, 648 experimental trials and 12 practice trials were processed.

#### 2.1.4. Design and Data Analysis

The experiment had a  $2 \times 2$  factorial design, resulting from the orthogonal combination of the within-subject factors repetition (repeated vs. novel) and ISI (constant vs. variable). The order of the ISI blocks (constant first vs. variable first) and the judgment-to-key assignment (left-shorter, right-longer vs. left-longer, right-shorter) were counterbalanced across participants.

Logistic functions were fitted to the data of each condition and for each participant. The point of subjective equality (PSE) was computed from each function as a measure of perceived duration. The PSE indicates the comparison duration which appears to be just as long as the standard. Larger PSEs indicate that the participant tends to perceive the comparison as shorter than the standard. In addition, DL was also computed from these functions as a measure of discrimination sensitivity. Larger DLs indicate poorer temporal discrimination.

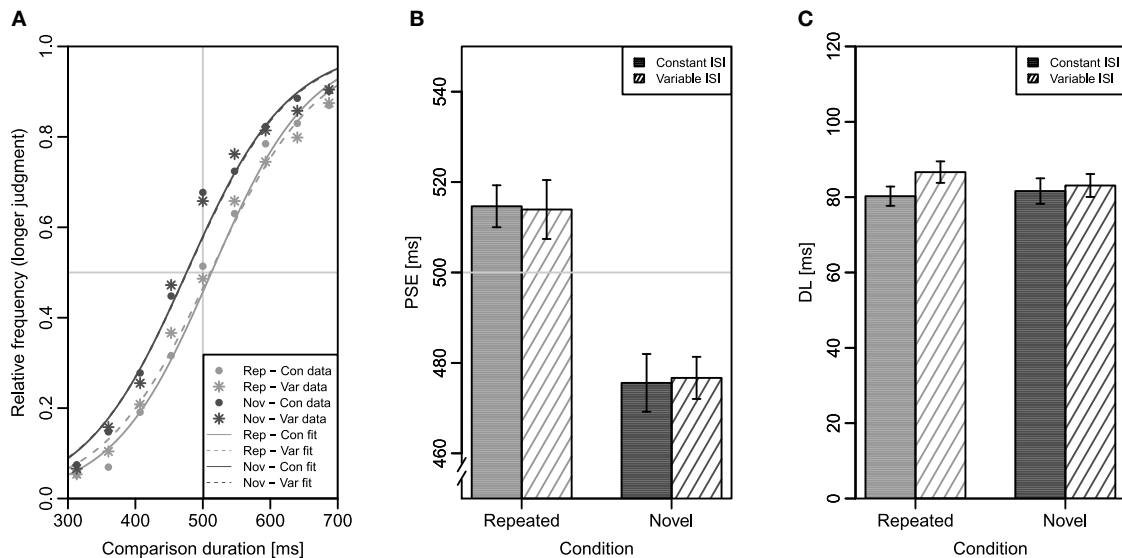
Finally, we analyzed response times (RTs) (see Birngruber et al., 2015, for another application of RT analyses in duration judgment tasks). First, we excluded all trials with RTs which were larger than 4000 ms because we considered them outliers (this led to the exclusion of 105 trials which is <0.6% of all trials). Then we computed mean RT as a function of comparison duration. Typically, RT in choice paradigms like the present shorter-longer-judgment task increases with discrimination difficulty (Birren and Botwinick, 1955; Sternberg, 1969). Thus, mean RT as a function of comparison duration should result in an inverted U-shaped function showing that participants need more time to decide whether the comparison was shorter or longer than the standard if the comparison duration is close to the PSE. To quantify the location of this inverted U-shaped function, we determined its first moment using the waveform moment analysis (Cacioppo and Dorfman, 1987; Ulrich et al., 1995). This location parameter represents the respective comparison duration at which the mean of the function is located<sup>1</sup>. This measure assesses the comparison duration that is most difficult to discriminate. We will, therefore, refer to this parameter as the point of maximal uncertainty (PMU). We determined PMU separately for each participant and each condition. A significance level of 0.05 was set for all significance tests and *p*-values were Greenhouse–Geisser corrected where appropriate.

#### 2.2. Results and Discussion

**Figure 2A** shows mean relative frequencies of “longer”-judgments as a function of comparison duration together with the fitted logistic function for each of the four conditions. Note that these functions are only for illustration, while the individual PSE and DL-values that entered the following analyses were derived from individually fitted psychometric functions.

A two-factor repeated measures ANOVA with the factors repetition (repeated vs. novel) and ISI (constant vs. variable) was conducted on PSE. **Figure 2B** depicts mean PSE as a function

<sup>1</sup>Consider  $m_i$  being the mean RT at comparison duration  $d_i$ ,  $i = 1, \dots, 9$ . First, these means are scaled as  $m_i^* = \frac{m_i}{\sum_{i=1}^9 m_i}$ ,  $i = 1, \dots, 9$ . Second, these scaled values are used to compute  $M = \sum_{i=1}^9 d_i \cdot m_i^*$ , that is, the location of the observed comparison duration-RT function.



**FIGURE 2 | Results of Experiment 1.** Error bars represent  $\pm 1$  standard error for within-subjects designs according to Morey (2008). **(A)** Mean relative frequencies of longer judgments as a function of comparison duration, in the four conditions (symbols) and fitted logistic functions for all participants' data (lines). Note that this plot is only for illustration; individually fitted logistic functions for each participant were the basis of

the statistical analyses. The horizontal light gray line indicates the 50% point; the vertical light gray line indicates the standard duration of 500 ms. **(B)** Mean point of subjective equality (PSE) in the four conditions. **(C)** Mean difference limen (DL) in the four conditions. Rep, repeated condition; Nov, novel condition; Con, constant inter-stimulus interval; Var, variable inter-stimulus interval.

of the two factors. A significant main effect of repetition was present,  $F_{(1, 31)} = 23.60$ ,  $MSE = 1975$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.43$ , indicating that mean PSE for repeated comparisons (514 ms) was larger than mean PSE for novel comparisons (476 ms). The main effect of ISI was not significant,  $F < 1$ , because PSE was identical for constant and variable ISIs (495 ms). The ANOVA revealed no interaction of the two factors,  $F < 1^2$ .

A two-factor repeated measures ANOVA with the same two factors was performed on DL. **Figure 2C** depicts mean DL. Neither the main effect of repetition,  $F < 1$ , nor the main effect of ISI,  $F_{(1, 31)} = 1.27$ ,  $MSE = 391.0$ ,  $p = 0.269$ ,  $\eta_p^2 = 0.04$ , were significant, indicating that discrimination sensitivity was almost identical for repeated and novel comparisons (83 and 82 ms) and for constant and variable ISIs (81 and 85 ms). The interaction of the factors was not significant either,  $F_{(1, 31)} = 1.12$ ,  $MSE = 171.5$ ,  $p = 0.298$ ,  $\eta_p^2 = 0.03$ .

To analyze RT, we conducted a three-factor repeated measures ANOVA with the factors repetition (repeated vs. novel), ISI (constant vs. variable), and comparison duration (nine levels,

313–687 ms) on mean RT. Mean RT as a function of comparison duration is depicted in **Figure 3A**. The main effect of repetition was significant,  $F_{(1, 31)} = 15.26$ ,  $MSE = 23,678$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.33$ , showing that RTs were generally longer for novel comparisons (637 ms) than for repeated ones (595 ms). As expected, RTs also changed across comparison durations causing a significant main effect of comparison duration,  $F_{(8, 248)} = 26.96$ ,  $MSE = 24,112$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.47$ , and a significant interaction of repetition and comparison duration,  $F_{(8, 248)} = 2.26$ ,  $MSE = 10,772$ ,  $p = 0.044$ ,  $\eta_p^2 = 0.07$ . No other effects of this ANOVA reached significance.

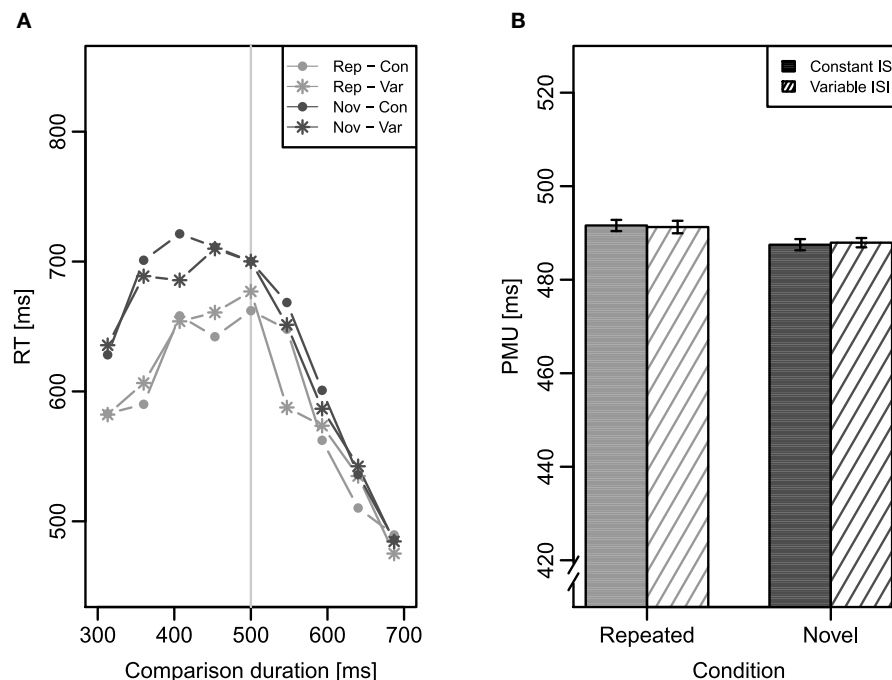
The PMUs which were calculated individually for each participant, served as the dependent variable of a two-factor repeated measures ANOVA with the factors repetition and ISI. Mean PMU for the four conditions can be found in **Figure 3B**. The significant main effect of repetition,  $F_{(1, 31)} = 9.12$ ,  $MSE = 48.80$ ,  $p = 0.005$ ,  $\eta_p^2 = 0.23$ , confirmed that the comparison duration-RT functions were slightly shifted, as expected. The mean PMU for the repeated condition (491 ms) was slightly larger than the mean PMU for the novel condition (488 ms). Neither a main effect of ISI nor an interaction of the two factors were evident (both  $F$ s  $< 1$ )<sup>3</sup>.

Taken together, the results of Experiment 1 show that novel stimuli are estimated to be longer than repeated stimuli of the same physical duration. This result was further supported by

<sup>2</sup>As mentioned before, the multiple presentation of the pseudowords throughout the experiment should not have influenced the effect of repetition condition on PSE because this effect is assumed to be short-lived. Nevertheless, we performed an additional repeated measures ANOVA with the factors repetition and experimental half (first half vs. second half) on PSE to check whether the repetition effect changed over time and hence with increasing number of stimulus presentations. The main effect of condition was identical to the one in the main analysis,  $F_{(1, 31)} = 23.60$ ,  $MSE = 1975$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.43$ . Although there was a significant main effect of experimental half on PSE,  $F_{(1, 31)} = 8.10$ ,  $MSE = 593$ ,  $p = 0.008$ ,  $\eta_p^2 = 0.21$  (first half: 501 ms, second half: 489 ms), the repetition effect was of similar size in the first (36 ms) and in the second half of the experiment (41 ms),  $F < 1$ .

<sup>3</sup>Additionally, we used an alternative method to estimate PMU. We fitted second degree polynomials to the comparison duration-RT functions of each condition and for each participant and determined their maxima. The comparison duration at which the maximum was located served as  $PMU_{poly}$ . Qualitatively, we observed the same pattern of results for  $PMU_{poly}$  as for PMU.





**FIGURE 3 | Results of Experiment 1.** Error bars represent  $\pm 1$  standard error for within-subjects designs according to Morey (2008). **(A)** Mean response time (RT) as a function of comparison duration in the four conditions. The vertical light gray line indicates

the standard duration of 500 ms. **(B)** Mean point of maximal uncertainty (PMU) in the four conditions. Rep, repeated condition; Nov, novel condition; Con, constant inter-stimulus interval; Var, variable inter-stimulus interval.

the RT results which showed that longest RT and thus the greatest uncertainty was observed at slightly shorter comparison durations for novel stimuli than for repeated stimuli. Participants' discrimination sensitivity was neither influenced by stimulus repetition nor by the ISI manipulation.

### 3. Experiment 2

The pseudowords in Experiment 1 did not convey high-level semantic information. Nevertheless, they followed the rules of German orthography and phonology and thus were pronounceable. It is therefore conceivable that the standards were retained in memory by subvocal rehearsal (Baddeley, 2012). This may have influenced the judged duration of repeated as compared to novel comparisons. Former research has shown that "illegal nonwords" like strings of consonants are more difficult to remember (Bowers, 1994) and harder to subvocalize than pseudowords (McCusker et al., 1981). We therefore conducted another experiment in which unpronounceable strings of consonants were used as stimuli.

#### 3.1. Method

##### 3.1.1. Participants

A fresh sample of 32 volunteers (19 female, 29 right-handed), aged between 18 and 33 years ( $M = 24.4$  years) participated in the experiment. All participants had normal or corrected-to-normal vision, and received course credit or €6. The

data of six additional participants was collected but had to be excluded from analyses due to the exclusion criteria already used in Experiment 1. All participants gave informed consent.

##### 3.1.2. Apparatus and Stimuli

Experiment 2 was identical to Experiment 1 except for the following changes. A MAC computer controlled stimulus presentation and recorded the participants' responses. The same VGA monitor was used as in Experiment 1. To generate a set of unpronounceable consonant strings, we transformed the set of stimuli from Experiment 1 as follows: Vowels in the pseudowords were replaced by consonants ("Y" was not used as replacement because it is sometimes pronounced like an "I" in German) whereby identical vowels in one pseudoword were replaced by the same consonant and different vowels were replaced by different consonants (e.g., MEILEG was changed to MKPLKG). The assignment from vowel to consonant was randomized for each word (e.g., MEILEG to MKPLKG and SEBSER to SMBSMR). In the end, a set of 104 consonant strings was created for each participant from which 81 were randomly selected to appear in the experiment. The items for the practice block were chosen from the remaining 23 strings of consonants.

##### 3.1.3. Procedure, Design, and Data Analysis

Again, RTs larger than 4000 ms were excluded (146 trials, <0.8% of all trials).

### 3.2. Results and Discussion

**Figure 4A** shows mean relative frequencies of “longer”-judgments and the fitted logistic functions for all four conditions. **Figure 4B** depicts mean PSE as a function of the factors repetition and ISI. As in Experiment 1, a significant main effect of repetition was obtained,  $F_{(1, 31)} = 18.90$ ,  $MSE = 4505$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.38$ , indicating again that the mean PSE for repeated comparisons (533 ms) was larger than the mean PSE for novel comparisons (481 ms). As before, the main effect of ISI was not significant,  $F < 1$ ; mean PSE was 504 ms for constant ISIs and 509 ms for variable ISIs. Like in Experiment 1, the ANOVA revealed no interaction of the two factors,  $F < 1^4$ .

Mean DL for the four conditions is shown in **Figure 4C**. The main effect of repetition on DL was significant in this experiment,  $F_{(1, 31)} = 9.70$ ,  $MSE = 204.6$ ,  $p = 0.004$ ,  $\eta_p^2 = 0.24$ , indicating that mean DL was larger for repeated trials (85 ms) than for novel trials (77 ms). There was a marginally significant main effect of ISI,  $F_{(1, 31)} = 3.79$ ,  $MSE = 618.2$ ,  $p = 0.061$ ,  $\eta_p^2 = 0.11$ , reflecting a trend for slightly better discrimination sensitivity when ISIs were constant (77 ms) than when they were variable (85 ms). The interaction of both factors was non-significant,  $F < 1$ .

Mean RT as a function of repetition condition and comparison duration is illustrated in **Figure 5A**. Mean RT was again longer for novel (666 ms) than for repeated comparisons (641 ms) resulting in a significant main effect of repetition,  $F_{(1, 31)} = 9.12$ ,  $MSE = 19,520$ ,  $p = 0.005$ ,  $\eta_p^2 = 0.23$ . As before, no main effect of ISI was observed,  $F < 1$ , with a mean RT of 663 ms in the constant ISI condition and a mean RT of 666 ms in the variable ISI condition. The main effect of comparison duration was significant,  $F_{(8, 248)} = 23.11$ ,  $MSE = 29,349$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.43$ , illustrating the typical reversed U-shaped comparison duration-RT function. As in Experiment 1, the factors repetition and comparison duration interacted significantly,  $F_{(8, 248)} = 3.28$ ,  $MSE = 9737$ ,  $p = 0.005$ ,  $\eta_p^2 = 0.10$ . No other effect was significant.

Mean PMU of the four conditions is depicted in **Figure 5B**. The ANOVA on PMU revealed again a significant main effect of repetition,  $F_{(1, 31)} = 12.82$ ,  $MSE = 51.20$ ,  $p = 0.001$ ,  $\eta_p^2 = 0.29$ , confirming that mean PMU was slightly larger for the repeated condition (493 ms) than for the novel condition (489 ms). Neither a main effect of ISI,  $F < 1$ , nor an interaction of the two factors was evident,  $F_{(1, 31)} = 2.72$ ,  $MSE = 22.39$ ,  $p = 0.109$ ,  $\eta_p^2 = 0.08^5$ .

Taken together, the results of Experiment 2 confirm the main finding of Experiment 1, namely that novel stimuli are overestimated in duration as compared to repeated stimuli. The

PMU analysis further supported this finding by showing that the comparison durations causing the longest RTs were smaller for novel than for repeated strings of consonants, meaning that the largest uncertainty was observed at shorter comparison durations for novel than for repeated stimuli. In contrast to Experiment 1, discrimination sensitivity benefitted from novel comparisons and marginally from a fixed temporal structure (ISI).

### 4. General Discussion

Two experiments were conducted in order to examine the influence of repetition on duration estimation with nonword stimuli. To this end, two nonwords were presented consecutively on each trial and the participants' task was to judge whether the second nonword (comparison) was shorter or longer than the first nonword (standard). Crucially, the comparison could either be a repetition of the standard (repeated comparison) or a different nonword (novel comparison). In Experiment 1, semantically low-associative pseudowords served as stimuli. In Experiment 2, unpronounceable strings of consonants were used as stimuli to test whether the results for pseudowords would transfer to stimulus material for which subvocalizing was not possible.

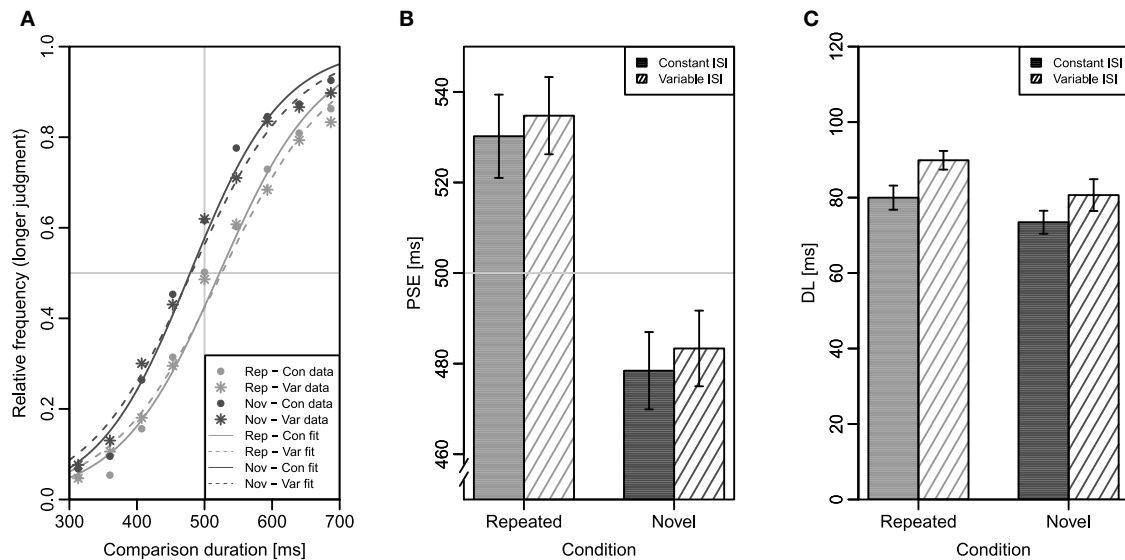
The results of the two experiments showed that the duration of the comparison was judged to be shorter for repeated than for novel stimuli, thereby replicating and extending the findings of Matthews (2011) who used a similar paradigm with more complex stimulus material (namely photographs of natural or social scenes, objects, and buildings). In the study by Matthews (2011), the effect might have been based on the repetition (for repeated comparisons) or change (for novel comparisons) of multiple low-level and high-level features. In a more recent study (Matthews, 2015), the effect was replicated for abstract stimuli, but these stimuli still contained a variety of different low-level features (like different colors and contrasts) and were constructed from various line drawings. The present study shows that the repetition or change of a simple, meaningless letter string was sufficient to generate the effect. The nonwords we used in the present experiments were well-controlled in low-level features (constant length, white font on black background), composed of familiar features (i.e., letters), and contained no obvious semantic information. Nevertheless, repeating these simple stimuli resulted in shorter judged durations as compared to presenting a different nonword as comparison. Thus, the difference in duration judgments for repeated and novel stimuli seems to be independent of the information complexity the stimuli contain.

Furthermore, immediate stimulus repetition influenced duration judgments irrespective of whether the nonwords were pronounceable (pseudowords) or unpronounceable (strings of consonants). The very similar results of Experiments 1 and 2 illustrate that the possibility to subvocalize the pseudowords and therefore to potentially experience increased processing fluency for repeated pseudowords (Johnston et al., 1985) is not crucial for duration judgment.

These consistent duration judgment results were further supported by RT analyses. It is assumed that participants

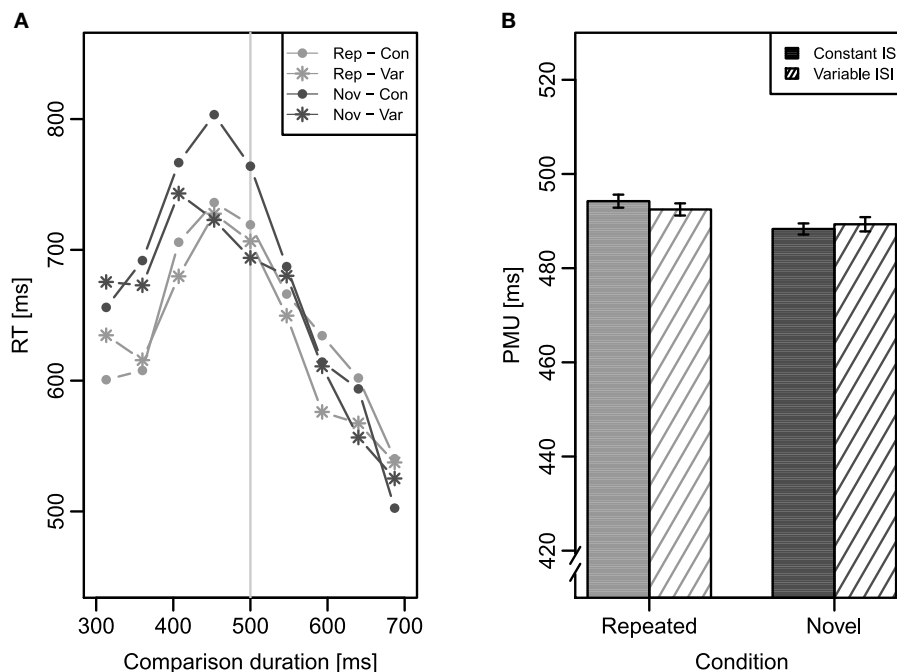
<sup>4</sup>Again an additional repeated measures ANOVA with the factors repetition and experimental half was conducted on PSE to check whether the repetition effect changed over the experimental course. The main effect of condition was identical to the one in the main analysis,  $F_{(1, 31)} = 18.90$ ,  $MSE = 4505$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.38$ . A significant main effect of experimental half,  $F_{(1, 31)} = 17.21$ ,  $MSE = 1530$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.36$ , demonstrated that PSE decreased from the first (521 ms) to the second half (481 ms) of the experiment. But importantly, the repetition effect was of almost equal size in the first (53 ms) and in the second half (51 ms) of the experiment,  $F < 1$ .

<sup>5</sup>As in Experiment 1, the additional analysis of  $PMU_{poly}$  yielded qualitatively the same pattern of results.



**FIGURE 4 | Results of Experiment 2.** Error bars represent  $\pm 1$  standard error for within-subjects designs according to Morey (2008). **(A)** Mean relative frequencies of longer judgments as a function of comparison duration in the four conditions (symbols) and fitted logistic functions for all participants' data (lines). Note that this plot is only for illustration; individually fitted logistic functions were the basis of the

statistical analyses. The horizontal light gray line indicates the 50% point; the vertical light gray line indicates the standard duration of 500 ms. **(B)** Mean point of subjective equality (PSE) in the four conditions. **(C)** Mean difference limen (DL) in the four conditions. Rep, repeated condition; Nov, novel condition; Con, constant inter-stimulus interval; Var, variable inter-stimulus interval.



**FIGURE 5 | Results of Experiment 2.** Error bars represent  $\pm 1$  standard error for within-subjects designs according to Morey (2008). **(A)** Mean response time (RT) as a function of comparison duration, for the four conditions. The vertical light gray line

indicates the standard duration of 500 ms. **(B)** Mean point of maximal uncertainty (PMU) for the four conditions. Rep, repeated condition; Nov, novel condition; Con, constant inter-stimulus interval; Var, variable inter-stimulus interval.

generally respond faster the more certain they are about a discrimination judgment (Birren and Botwinick, 1955; Sternberg, 1969). Accordingly, comparison durations which are subjectively similar to the standard duration should result in the longest RTs. To determine whether these points of maximal uncertainty were shifted against each other depending on whether comparisons were repeated or novel, we used the waveform moment analysis (Cacioppo and Dorfman, 1987) to determine the means of the comparison duration-RT functions. Indeed, the calculated PMUs were significantly smaller for novel than for repeated comparisons. Thus, participants experienced shorter durations as equivalent to the standard duration when the comparison was novel than when the comparison was repeated. Hence, the PMU analyses complemented the duration judgment results and corroborated that repeated comparisons were perceived as being shorter than novel comparisons.

Moreover, RTs were generally longer for novel than for repeated comparisons. We assume that this main effect does not necessarily reflect lower decision certainty for novel than for repeated trials because discrimination sensitivity (DL) was either equal in the two conditions (Experiment 1) or even superior for novel comparisons (Experiment 2). This finding might rather reflect general processing advantages for repeated stimuli (Pashler and Baylis, 1991; Bentin and McCarthy, 1994) which could be independent from the quality of temporal discrimination. It should be noted, however, that RT advantages for repeated stimuli are usually reported for tasks in which participants have to judge the identity of a stimulus, whereas in our case the identity of the stimulus is actually irrelevant for the response. Therefore, it is not entirely clear, whether the same mechanism are at work both in stimulus discrimination and duration discrimination when stimuli are repeated.

Additionally we were interested in whether a predictable time course within each trial (i.e., a constant ISI between the standard and the comparison) would improve the participants' temporal discrimination sensitivity. This was suggested by Tse et al. (2004) who argued that a constant ISI might induce rhythm perception which could alter duration judgments. To investigate this issue, a constant ISI was used in one half of both experiments while ISIs varied slightly in the other half of both experiments. Only in Experiment 2 did participants show a tendency to discriminate durations better when a predictable time course was used. By and large, however, manipulation of ISI had little if any effect in the present experiments. It could be speculated that a rhythm might only be induced by stream-based paradigms in which a series of standards is presented prior to the comparison. Alternatively, one could argue that the variable ISIs in the present experiments only varied slightly (247–380 ms) in duration; a range which might not have disrupted the anticipation of the comparison onset enough. It is also conceivable that rhythm effects are generally

more pronounced for empty time intervals which are defined by two markers at the beginning and the end of a stimulus than for filled intervals which were used in the present study.

Generally, it is well-known in experimental psychology that the repetition of stimuli influences their cognitive processing. The repetition of stimuli has often been understood as a special case of priming (Henson, 2003; Schacter et al., 2007) whereby repetition priming has shown to decrease reaction times and increase discrimination performance for repeated stimuli (Bentin and McCarthy, 1994). The neural mechanism of “repetition suppression,” which describes a reduced neural response to repeated stimulus presentation, has been argued to be responsible for at least some of the behavioral effects of stimulus repetition (Wig et al., 2005). Since the size of the neural response has also been suggested to form the basis of duration perception (Pariyadath and Eagleman, 2012), the present results might be linked to repetition suppression.

Interestingly, there is evidence that this mechanism is not necessarily limited to one-on-one repetitions following an all-or-nothing principle, but that gradual differences between standards and comparisons also shape duration judgments. Specifically, the size of the oddball effect increases the more the comparisons deviate from the standards (Schindel et al., 2011; Pariyadath and Eagleman, 2012). Furthermore, it has been shown that not only repetition but also expectation can influence duration judgments. For example, Pariyadath and Eagleman (2007) reported that the duration of a number embedded in a predictable sequence (e.g., 1 2 3 4) was judged as similarly long as a number embedded in a sequence of repeated numbers (e.g., 1 1 1 1) but as shorter than a number embedded in an unpredictable sequence (e.g., 1 3 5 2). Recent results by Matthews (2015) suggest that the effects of repetition and expectation on duration judgment interact in a complex manner. Surprisingly, Matthews reported smaller repetition effects for frequent than for infrequent repetitions when the likelihood of stimulus repetition was manipulated. Future research needs to further disentangle the effects of repetition on the one hand and expectation on the other hand.

The present study replicates and extends previous findings concerning the effect of immediate stimulus repetition on duration perception. Furthermore, the results clearly suggest that changes of simple, meaningless stimuli with similar low-level features are sufficient to induce a shorter perceived duration of repetitions, and that the temporal structure within the reminder task has no pronounced effect on duration judgments.

## Acknowledgments

This research was supported by a grant of the Deutsche Forschungsgemeinschaft (UL 116/10-2) and the Open Access Publishing Fund of the University of Tübingen.

## References

- Baddeley, A. (2012). Working memory: theories, models, and controversies. *Annu. Rev. Psychol.* 63, 1–29. doi: 10.1146/annurev-psych-1207-10-100422
- Bentin, S., and McCarthy, G. (1994). The effects of immediate stimulus repetition on reaction time and event-related potentials in tasks of different complexity. *J. Exp. Psychol.* 20, 130–149.
- Birngruber, T., Schröter, H., and Ulrich, R. (2014). Duration perception of visual and auditory oddball stimuli: does judgment task modulate the temporal



- oddball effect? *Attent. Percept. Psychophys.* 76, 814–828. doi: 10.3758/s13414-013-0602-2
- Birngruber, T., Schröter, H., and Ulrich, R. (2015). Introducing a control condition in the classic oddball paradigm: oddballs are overestimated in duration not only because of their oddness. *Attent. Percept. Psychophys.* 77, 1737–1749. doi: 10.3758/s13414-015-0868-7
- Birren, J. E., and Botwinick, J. (1955). Speed of response as a function of perceptual difficulty and age. *J. Gerontol.* 10, 433–436.
- Bowers, J. S. (1994). Does implicit memory extend to legal and illegal nonwords? *J. Exp. Psychol.* 20, 534–549.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spat. Vis.* 10, 433–436.
- Cacioppo, J. T., and Dorfman, D. D. (1987). Wave-form moment analysis in psychophysiological research. *Psychol. Bull.* 102, 421–438.
- Chen, K.-M., and Yeh, S.-L. (2009). Asymmetric cross-modal effects in time perception. *Acta Psychol.* 130, 225–234. doi: 10.1016/j.actpsy.2008.12.008
- Droit-Volet, S., Fayolle, S. L., and Gil, S. (2011). Emotion and time perception: effects of film-induced mood. *Front. Integr. Neurosci.* 5:33. doi: 10.3389/fnint.2011.00033
- Eagleman, D. M. (2008). Human time perception and its illusions. *Curr. Opin. Neurobiol.* 18, 131–136. doi: 10.1016/j.conb.2008.06.002
- Goldstone, S., and Lhamon, W. T. (1974). Studies of auditory-visual differences in human time judgment: 1. Sounds are judged longer than lights. *Percept. Mot. Skills* 39, 63–82.
- Grondin, S. (2010). Timing and time perception: a review of recent behavioral and neuroscience findings and theoretical directions. *Attent. Percept. Psychophys.* 72, 561–582. doi: 10.3758/APP.72.3.561
- Henson, R. N. A. (2003). Neuroimaging studies of priming. *Prog. Neurobiol.* 70, 53–81. doi: 10.1016/S0301-0082(03)00086-8
- Johnston, W. A., Dark, V. J., and Jacoby, L. L. (1985). Perceptual fluency and recognition judgments. *J. Exp. Psychol.* 11, 3–11.
- Kim, E., and McAuley, J. D. (2013). Effects of pitch distance and likelihood on the perceived duration of deviant auditory events. *Attent. Percept. Psychophys.* 75, 1547–1558. doi: 10.3758/s13414-013-0490-5
- Matthews, W. J. (2011). Stimulus repetition and the perception of time: the effects of prior exposure on temporal discrimination, judgment, and production. *PLoS ONE* 6:e19815. doi: 10.1371/journal.pone.0019815
- Matthews, W. J. (2015). Time perception: the surprising effects of surprising stimuli. *J. Exp. Psychol.* 144, 172–197. doi: 10.1037/xge0000041
- Matthews, W. J., Stewart, N., and Wearden, J. H. (2011). Stimulus intensity and the perception of duration. *J. Exp. Psychol.* 37, 303. doi: 10.1037/a0019961
- McCusker, L. X., Hillinger, M. L., and Bias, R. G. (1981). Phonological recoding and reading. *Psychol. Bull.* 89, 217–245.
- Morey, R. D. (2008). Confidence intervals from normalized data: a correction to Cousineau (2005). *Reason* 4, 61–64.
- New, J. J., and Scholl, B. J. (2009). Subjective time dilation: spatially local, object-based, or a global visual experience? *J. Vis.* 9, 1–11. doi: 10.1167/9.2.4
- Pariyadath, V., and Eagleman, D. (2007). The effect of predictability on subjective duration. *PLoS ONE* 2:e1264. doi: 10.1371/journal.pone.0001264
- Pariyadath, V., and Eagleman, D. M. (2012). Subjective duration distortions mirror neural repetition suppression. *PLoS ONE* 7:e49362. doi: 10.1371/journal.pone.0049362
- Pashler, H., and Baylis, G. C. (1991). Procedural learning: II. Intertrial repetition effects in speeded-choice tasks. *J. Exp. Psychol.* 17, 33–48. doi: 10.1037/0278-7393.17.1.33
- Pelli, D. G. (1997). The videotoolbox software for visual psychophysics: transforming numbers into movies. *Spat. Vis.* 10, 437–442. doi: 10.1163/156856897X00366
- Rammesayer, T. H., and Verner, M. (2015). Larger visual stimuli are perceived to last longer from time to time: the internal clock is not affected by nontemporal visual stimulus size. *J. Vis.* 15, 1–11. doi: 10.1167/15.3.5
- Schacter, D. L., Wig, G. S., and Stevens, W. D. (2007). Reductions in cortical activity during priming. *Curr. Opin. Neurobiol.* 17, 171–176. doi: 10.1016/j.conb.2007.02.001
- Schindel, R., Rowlands, J., and Arnold, D. H. (2011). The oddball effect: perceived duration and predictive coding. *J. Vis.* 11, 1–9. doi: 10.1167/11.2.17
- Sternberg, S. (1969). The discovery of processing stages: extensions of Donders' method. *Acta Psychol.* 30, 276–315. doi: 10.1016/0001-6918(69)90055-9
- Sturm, W., and Willmes, K. (1999). *Verbal und Nonverbaler Lerntest (VLT/NVLT)*. Göttingen: Hogrefe.
- Thomas, E. A., and Cantor, N. E. (1976). Simultaneous time and size perception. *Percept. Psychophys.* 19, 353–360. doi: 10.3758/BF03204243
- Tse, P. U., Intriligator, J., Rivest, J., and Cavanagh, P. (2004). Attention and the subjective expansion of time. *Percept. Psychophys.* 66, 1171–1189. doi: 10.3758/BF03196844
- Ulrich, R., Nitschke, J., and Rammesayer, T. (2006). Perceived duration of expected and unexpected stimuli. *Psychol. Res.* 70, 77–87. doi: 10.1007/s00426-004-0195-4
- Ulrich, R., Wing, A. M., and Rinkenauer, G. (1995). Amplitude and duration scaling of brief isometric force pulses. *J. Exp. Psychol.* 21, 1457–1472. doi: 10.1037/0096-1523.21.6.1457
- Wearden, J. H., Edwards, H., Fakhri, M., and Percival, A. (1998). Why “sounds are judged longer than lights”: application of a model of the internal clock in humans. *Q. J. Exp. Psychol.* 51, 97–120.
- Wig, G. S., Grafton, S. T., Demos, K. E., and Kelley, W. M. (2005). Reductions in neural activity underlie behavioral components of repetition priming. *Nat. Neurosci.* 8, 1228–1233. doi: 10.1038/nn1515

**Conflict of Interest Statement:** 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.

Copyright © 2015 Birngruber, Schröter and Ulrich. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Visual-auditory differences in duration discrimination of intervals in the subsecond and second range

Thomas H. Rammsayer<sup>1,2\*</sup>, Natalie Borter<sup>1,2</sup> and Stefan J. Troche<sup>3</sup>

<sup>1</sup> Institute of Psychology, University of Bern, Bern, Switzerland, <sup>2</sup> Center for Cognition, Learning, and Memory, University of Bern, Bern, Switzerland, <sup>3</sup> Department of Psychology and Psychotherapy, University of Witten/Herdecke, Witten, Germany

## OPEN ACCESS

### Edited by:

Lihan Chen,  
Peking University, China

### Reviewed by:

Matthew S. Matell,  
Villanova University, USA  
Hugo Merchant,  
Universidad Nacional Autónoma  
de México, Mexico

### \*Correspondence:

Thomas H. Rammsayer  
thomas.rammsayer@psy.unibe.ch

### Specialty section:

This article was submitted to  
Perception Science,  
a section of the journal  
Frontiers in Psychology

**Received:** 05 March 2015

**Accepted:** 08 October 2015

**Published:** 26 October 2015

### Citation:

Rammsayer TH, Borter N  
and Troche SJ (2015) Visual-auditory  
differences in duration discrimination  
of intervals in the subsecond  
and second range.  
Front. Psychol. 6:1626.  
doi: 10.3389/fpsyg.2015.01626

A common finding in time psychophysics is that temporal acuity is much better for auditory than for visual stimuli. The present study aimed to examine modality-specific differences in duration discrimination within the conceptual framework of the Distinct Timing Hypothesis. This theoretical account proposes that durations in the lower milliseconds range are processed automatically while longer durations are processed by a cognitive mechanism. A sample of 46 participants performed two auditory and visual duration discrimination tasks with extremely brief (50-ms standard duration) and longer (1000-ms standard duration) intervals. Better discrimination performance for auditory compared to visual intervals could be established for extremely brief and longer intervals. However, when performance on duration discrimination of longer intervals in the 1-s range was controlled for modality-specific input from the sensory-automatic timing mechanism, the visual-auditory difference disappeared completely as indicated by virtually identical Weber fractions for both sensory modalities. These findings support the idea of a sensory-automatic mechanism underlying the observed visual-auditory differences in duration discrimination of extremely brief intervals in the millisecond range and longer intervals in the 1-s range. Our data are consistent with the notion of a gradual transition from a purely modality-specific, sensory-automatic to a more cognitive, amodal timing mechanism. Within this transition zone, both mechanisms appear to operate simultaneously but the influence of the sensory-automatic timing mechanism is expected to continuously decrease with increasing interval duration.

**Keywords:** duration discrimination, sensory modality, subsecond range, second range, distinct timing hypothesis, common timing hypothesis, timing mechanisms

## INTRODUCTION

A common finding in time psychophysics is that temporal acuity is much better for auditorily than for visually presented stimuli (Penney and Tourret, 2005; van Wassenhove, 2009; Merchant et al., 2015). This also applies to perceived duration and duration discrimination as two aspects of interval timing. Perceived duration reflects the subjectively experienced duration of a given stimulus interval, while duration discrimination refers to the ability to discriminate the smallest possible difference in duration between two temporal intervals. A large number of studies demonstrated that, when a visual and an auditory stimulus are presented for the same physical time, the perceived duration of the auditory stimulus is longer than the perceived duration of the visual

one (e.g., Goldstone and Lhamon, 1974; Walker and Scott, 1981; Wearden et al., 1998; Penney et al., 2000; Penney, 2003; Droit-Volet et al., 2004; Ortega et al., 2009). With regard to duration discrimination, the available data indicate better temporal discrimination of auditory compared to visually presented intervals (for concise reviews see Grondin, 2003; Rammsayer, 2014). A main objective of the present study was to contribute to a better understanding of the mechanisms involved in visual-auditory differences in temporal discrimination of extremely brief intervals in the range of 10s of milliseconds and longer intervals in the 1-s range.

There are two major conceptual frameworks to account for the timing of extremely brief and longer intervals: the *Common Timing Hypothesis* and the *Distinct Timing Hypothesis* (cf. Rammsayer and Troche, 2014). Broadly speaking, the Common Timing Hypothesis assumes a single, unitary timing mechanism irrespective of interval duration, whereas the Distinct Timing Hypothesis proposes two dissociable mechanisms for the timing of durations in the sub-second and second range, respectively.

The first psychophysical models of interval timing in the subsecond and second range, developed by Creelman (1962) and Treisman (1963), proposed a common timing mechanism based on neural counting. According to these models, a neural pacemaker generates pulses and the number of pulses associated with a physical time interval constitutes the internal time code of this interval. Thus, the higher the pulse rate, the better the temporal resolution of the timing mechanism will be, which is functionally equivalent to better performance on interval timing. More recent theoretical accounts of interval timing, the most well-known of which is Scalar Timing Theory (e.g., Gibbon and Church, 1984; Church, 2003; Allman et al., 2014), also assume such a unitary timing mechanism (Killeen and Weiss, 1987; Rammsayer and Ulrich, 2001; Grondin, 2010). Although direct experimental evidence for the notion of a single timing mechanism underlying duration discrimination in the subsecond and second range is difficult to obtain, some indirect evidence can be derived from the failure to detect a decrease in precision across two ranges of interval duration due to the breakpoint of an interval timing mechanism (Lewis and Miall, 2009). Such break points are to be expected if distinct timing mechanisms, with various levels of absolute precision, were used for measuring intervals of different durations (Rammsayer, 1996; Gibbon et al., 1997; Grondin, 2014).

Most likely, Münsterberg (1889) was the first to propose two distinct timing mechanisms underlying interval timing in the subsecond and second range, respectively. He assumed that durations less than approximately 300 ms can be perceived directly, whereas longer durations need to be formed by higher mental processes. Similarly, Michon (1985) put forward the idea that temporal processing of intervals longer than approximately 500 ms is cognitively mediated, whereas temporal processing of shorter intervals is “of a highly perceptual nature, fast, parallel and not accessible to cognitive control” (Michon, 1985, p. 40).

The Distinct Timing Hypothesis is supported by several studies (e.g., Rammsayer and Lima, 1991; Rammsayer and Ulrich, 2011) employing a dual-task paradigm with a temporal

primary task (e.g., duration discrimination) and a secondary non-temporal cognitive task (e.g., word learning). In these studies, temporal discrimination of intervals ranging from 50 to 100 ms was not affected by the non-temporal secondary task, whereas discrimination of longer intervals in the 1-s range was markedly impaired by the same secondary task. These findings were consistent with Michon’s (1985) notion that temporal processing of extremely brief intervals can be regarded as sensory-automatic in nature and beyond cognitive control, while temporal processing of longer intervals demands cognitive resources. This pattern of results was corroborated by pharmacopsychological studies showing a differential effect of pharmacological agents on temporal discrimination as a function of interval duration. Drugs that interfere with working memory functioning, such as benzodiazepines, strongly impact performance on duration discrimination in the 1-s range without affecting the 10s-of-ms range (for a concise review see Rammsayer, 2008). Also findings from functional neuroimaging studies corroborate the concept of a sensory-automatic system for the timing of intervals in the range of 10s of milliseconds and a cognitively controlled, higher-order system for temporal processing of longer intervals (Lewis and Miall, 2003, 2006). Several more recent studies, also proceeding from Michon’s (1985) conception of two distinct timing mechanisms, implied that the transition from automatic-sensory to cognitively controlled timing lies closer to 250 ms than to 500 ms (e.g., Buonomano et al., 2009; Spencer et al., 2009).

Most studies on visual-auditory differences in duration discrimination, more or less implicitly, refer to the Common Timing Hypothesis to account for their findings (cf. Grondin, 2003; Rammsayer, 2014). Within this conceptual framework, better performance on auditory duration discrimination is generally ascribed to an increased number of pulses accumulated during a given time interval in the case of auditory compared to visual stimuli. This increased number of pulses yields finer temporal resolution and, thus, better timing accuracy for auditory compared to visual intervals.

Up to date and to the best of our knowledge, no experimental study appears to exist that directly addressed visual-auditory differences in duration discrimination against the theoretical background of the Distinct Timing Hypothesis. Therefore, the major goal of the present study was to explore whether there is evidence for the notion of different timing mechanisms underlying visual-auditory differences in duration discrimination of extremely brief intervals in the range of 10s of milliseconds and longer intervals in the 1-s range.

Our theoretical point of departure was provided by two recent studies applying a confirmatory factor analysis (CFA) approach. In their study on visual-auditory differences in temporal information processing, Stauffer et al. (2012) put forward the idea that modality-specific differences develop at the level of sensory-automatic processing, whereas higher-order cognitive temporal processing was assumed to be amodal and, thus, independent of sensory modality. This notion of an amodal mechanism for temporal processing of longer intervals is supported by the finding of similar tuning properties of neurons in the supplementary motor area to durations in the 450- to 1000-ms range across sensory modalities (Merchant et al., 2013a,b).

In another, more recent CFA study on the internal structure of auditory interval timing in the subsecond and 1-s range, Rammsayer and Troche (2014) concluded that the assumption of two distinct mechanisms underlying the processing of extremely brief and longer intervals might be more appropriate than the assumption of a unitary timing mechanism. Most importantly, however, for the 1-s range, they proposed a shared influence of the sensory-automatic and the cognitive timing mechanism. This shared influence originates from the notion of a transition zone from primarily sensory-automatic to primarily cognitive temporal processing (Hellström and Rammsayer, 2002; Buonomano et al., 2009; Rammsayer and Ulrich, 2011). Within this transition zone, there may be a substantial degree of sensory-automatic and cognitive processing overlap as both mechanisms operate simultaneously. Thus, temporal processing of longer intervals in the 1-s range is assumed to be controlled by and functionally related to both sensory-automatic and cognitive temporal processing.

In the present study, we transferred these conclusions to visual-auditory differences in duration discrimination of extremely brief and longer intervals. By doing so, we arrived at two predictions. First, if the timing of longer durations involves not only cognitive processes but also depends, at least to some degree, on input from the sensory-automatic timing system, then visual-auditory differences in duration discrimination observed with extremely brief intervals should also become evident for longer intervals. Our second prediction, therefore, was that visual-auditory differences observed with longer intervals in the 1-s range can be explained by modality-specific differences in initial sensory-automatic processing. More precisely, if performance on duration discrimination of longer intervals, in fact, depends on sensory-automatic as well as cognitive processes, then the relative contribution of the sensory-automatic mechanism should become evident when performance scores on auditory (visual) duration discrimination with longer intervals are statistically controlled for performance on auditory (visual) duration discrimination obtained for extremely brief intervals. With such a methodological approach, the visual-auditory difference should decrease, or even disappear, in the adjusted performance scores for longer intervals, if modality-specific differences in duration discrimination indeed originate from the sensory-automatic level of temporal information processing. This line of reasoning, underlying Predictions 1 and 2, implies the following two assumptions: (1) It is possible to dissociate the contribution of the temporal processing of extremely brief intervals from that associated with cognitive processing of longer intervals and (2) the sensory-automatic mechanism is independent of the cognitive timing mechanism.

To test our predictions, participants performed auditory and visual two-alternative forced-choice duration discrimination tasks with extremely brief intervals in the subsecond range and longer intervals in the 1-s range. The durations of the standard intervals were 50 ms for the extremely brief and 1000 ms for the longer intervals. These standard durations were chosen because the hypothetical shift from one timing mechanism to the other is supposed to occur somewhere between 100 and 500 ms (Michon, 1985; Buonomano et al., 2009; Spencer et al., 2009). Furthermore,

it should be noted that, when participants are required to judge the duration of time intervals, many of them use counting as a non-temporal auxiliary strategy. Because this auxiliary counting strategy becomes effective for measuring intervals longer than approximately 1200 ms (Grondin et al., 1999, 2004), the “long” standard duration was chosen not to exceed this critical value.

## MATERIALS AND METHODS

### Participants

Twenty male and 26 female undergraduate students participated in the present study. Participants' age ranged from 18 to 28 years (mean age  $\pm$  standard deviation:  $22.7 \pm 2.5$  years). All participants were naïve with regard to the purpose of the study and reported normal hearing and normal or corrected-to-normal vision. The study was approved by the ethics committee of the Faculty of Human Sciences, University of Bern, and all participants gave their written informed consent.

### Procedure

Temporal stimuli were auditory and visual intervals. Auditory stimuli were white-noise signals presented through headphones (Vivanco SR85) at an intensity of 63 dB(A) SPL. Visual stimuli were generated by a red LED (diameter:  $0.48^\circ$ , viewing distance: 60 cm, luminance:  $48 \text{ cd/m}^2$ ) positioned at eye level of the participant. Testing took place in a sound-attenuated room with constant ambient light.

Performance on interval timing for extremely brief and longer intervals was assessed by one block of auditory and one block of visual intervals for each time range. Each of these four blocks comprised 64 trials, and each trial consisted of a constant standard and a variable comparison interval presented with an interstimulus interval of 900 ms. The duration of the standard interval was 50 ms for the extremely brief intervals and 1000 ms for the longer ones. The duration of the comparison interval was varied according to the weighted up-down method (Kaernbach, 1991), an adaptive rule to estimate  $x_{.25}$  and  $x_{.75}$  of the psychometric function of each participant. With this psychophysical approach,  $x_{.25}$  and  $x_{.75}$  indicate the duration of the two comparison intervals at which the response “longer” was given with a probability of 0.25 and 0.75, respectively. Each experimental block consisted of two series of 32 trials converging to  $x_{.25}$  and  $x_{.75}$ , respectively. For each series, the presentation order of the standard and the comparison interval was randomized and balanced. That way, standard and comparison intervals were presented first in 50% of the trials. Trials from both series were randomly interleaved within a block.

To estimate  $x_{.25}$  for the extremely brief intervals, the comparison interval was increased for Trials 1–6 by 3 ms if the participant had judged the standard interval to be longer and decreased by 9 ms after a “short” response. For Trials 7–32, the duration of the comparison interval was increased by 2 ms and decreased by 6 ms, respectively. The opposite step sizes were employed for  $x_{.75}$ . The initial durations of the comparison interval were 15 ms below and above the standard interval for  $x_{.25}$  and  $x_{.75}$ , respectively. For the discrimination

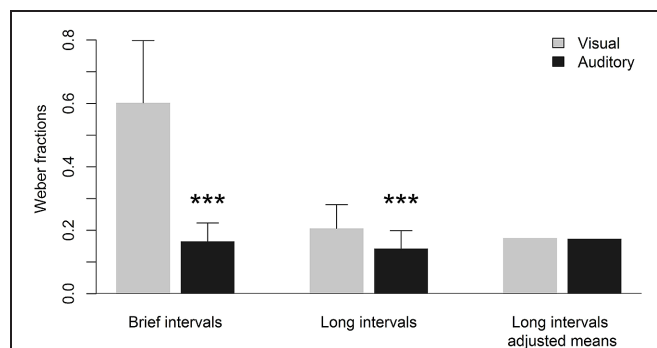


of longer intervals, the initial values of the comparison interval were 500 ms and 1,500 ms for  $x_{.25}$  and  $x_{.75}$ , respectively. To estimate  $x_{.25}$ , the duration of the comparison interval was increased by 100 ms if the standard interval was judged longer and decreased by 300 ms after a “short” response. For Trials 7–32, the duration of the comparison interval was increased by 25 ms and decreased by 75 ms, respectively. Again, the opposite step sizes were employed for  $x_{.75}$ .

Order of the four blocks was counterbalanced across participants. Prior to each block, practice trials were presented to familiarize participants with the task and to ensure that they understood the instructions. Participants were instructed to decide whether the first or the second interval was longer and to indicate their answers by pushing one of two designated response buttons. Each response was followed by visual correctness feedback presented on a monitor screen. As a psychophysical indicator of performance on duration discrimination, the difference limen (DL) was computed. Following Luce and Galanter (1963), DL was defined as half the interquartile range  $[(x_{.75} - x_{.25})/2]$ . With this performance measure, smaller DL values indicate better discrimination performance. More detailed information on our psychophysical approach can be found in Rammsayer (2012).

## RESULTS

Descriptive statistics of performance on duration discrimination as indicated by DL values are given in **Table 1**. For both extremely brief and longer intervals, smaller DL values and, thus, better performance on duration discrimination, were observed for auditory compared to visual stimuli. Subsequent  $t$ -tests revealed that these visual-auditory differences in DL values were statistically significant (see **Table 1**). In **Figure 1**, these visual-auditory differences are displayed graphically. For enhancing the presentation of results and to facilitate a comparison across the two ranges of interval duration, Weber fractions (DL/standard interval) are diagrammed instead of absolute DL values (cf. Killeen and Weiss, 1987; Rammsayer and Grondin, 2000). The outcome of these statistical analyses is not inconsistent with our first prediction. This prediction proceeded from the assumption that temporal processing of longer intervals not only involves cognitive processes but also depends, to some degree, on input from the sensory-automatic timing mechanism. In this case, a



**FIGURE 1 | Mean Weber fractions and standard deviations for visual and auditory duration discrimination of brief (50-ms standard duration) and longer (1000-ms standard duration) intervals as well as adjusted mean Weber fractions for duration discrimination of longer intervals.**

Adjusted means represent predicted mean Weber fractions controlled for the linear effect of sensory-automatic processing in the respective sensory modality. \*\*\* significantly different from respective visual duration discrimination ( $p < 0.001$ ).

visual-auditory difference in duration discrimination observed for extremely brief intervals in the 10s-of-ms range should also become evident for longer intervals in the 1-s range.

Next, we evaluated our second prediction assuming that the visual-auditory difference in duration discrimination of longer intervals is caused by the visual-auditory difference in duration discrimination observed for extremely brief intervals. In other words, we examined whether the visual-auditory difference of longer intervals depends on the input from the sensory-automatic timing system. If this prediction is true, the visual-auditory difference of longer intervals should disappear after statistical removal of the visual-auditory effect obtained with extremely brief intervals. To test this prediction, analysis of covariance (cf. Lee, 1975; Kirk, 1995; Tabachnick and Fidell, 2013) was applied.

In general terms, this statistical approach is an extension of analysis of variance as main effects and interactions are assessed after adjusting the dependent variables for the influence of at least one covariate for each dependent variable. Thus, analysis of covariance represents a combination of regression analysis and analysis of variance. In case of a within-subject design, separate regression analyses are used to adjust each dependent variable for the influence of at least one covariate. Then, in a second step, a repeated-measurement analysis of variance is performed on the adjusted values (Tabachnick and Fidell, 2013).

Following these considerations, a repeated-measures analysis of covariance was conducted with performance on duration discrimination with auditory and visual intervals in the 1-s range as dependent variables using BMDP 2V statistical software (Dixon, 1988). Again, for enhancing the presentation of results, Weber fractions were computed and analyzed. By applying analysis of covariance, each dependent variable (in the present case: performance on visual and auditory duration discrimination of longer intervals) was adjusted for the input from the modality-specific sensory-automatic timing system as reflected by performance on visual and auditory duration

**TABLE 1 | Mean difference limen (DL) values ( $M$ ) and standard deviations ( $SD$ ) in ms for visual and auditory duration discrimination of brief (50-ms standard duration) and longer (1000-ms standard duration) intervals.**

	Visual		Auditory		$t$	$d_z$
	$M$	$SD$	$M$	$SD$		
Brief intervals	30.1	9.8	8.3	2.9	19.22***	2.83
Longer intervals	206.4	74.8	141.5	56.6	5.78***	0.85

Also displayed are  $t$  values and Cohen's  $d_z$  as effect size estimate for modality-related differences ( $df = 45$ ;  $N = 46$ ).

\*\*\*  $p < 0.001$ .

discrimination of extremely brief intervals, respectively. This was achieved by two regressions. The first one regressed out visual duration discrimination of extremely brief intervals from visual duration discrimination of longer intervals, while the second one regressed out auditory duration discrimination of extremely brief intervals from auditory duration discrimination of longer intervals. The resulting adjusted means were evaluated by using the grand mean of the covariates as predictor for both regressions (for the regression equations see Kirk, 1995, p. 725). After adjusting the dependent variables for the visual-auditory difference resulting from the sensory-automatic timing system, the modality-related difference for duration discrimination of longer intervals disappeared,  $F(1,44) = 1.77$ ,  $p = 0.19$ . Performance on auditory and visual duration discrimination of longer intervals was virtually identical as indicated by adjusted mean Weber fractions of 0.175 and 0.173 for auditory and visual intervals, respectively. The adjusted means of longer intervals after controlling for the influence of the sensory-automatic timing system are depicted in **Figure 1**. As the adjusted means represent the predicted means for the visual and auditory sensory modality, respectively, there is no individual variability and, thus, no standard deviations can be reported. The outcome of the analysis of covariance provided clear evidence for the notion of a modality-specific sensory-automatic timing system. When the influence of this system was statistically controlled for, a visual-auditory difference in duration discrimination of longer intervals could no longer be established.

## DISCUSSION

The aim of the present study was to systematically investigate modality-specific differences in duration discrimination within the conceptual framework of the Distinct Timing Hypothesis. For this purpose, performance on duration discrimination of extremely brief and longer intervals in the auditory and visual modality was assessed by means of a within-subjects design. Proceeding from a modified version of the Distinct Timing Hypothesis, introduced by Rammsayer and Troche (2014), and from Stauffer et al.'s (2012) notion that modality-specific differences develop at the level of sensory-automatic processing rather than at the cognitive level, two predictions were made. First, if temporal processing of longer intervals in the 1-s range is, at least to some degree, dependent on input from the sensory-automatic timing mechanism, then a visual-auditory difference in duration discrimination observed for extremely brief intervals in the range of 10s of milliseconds should also become evident for longer intervals. Second, if the visual-auditory difference results from the sensory-automatic stage of temporal information processing, it should be reduced for duration discrimination of longer intervals after statistical removal of the visual-auditory effect originating from the sensory-automatic timing mechanism. Both these predictions were confirmed in the present study: superior discrimination performance for auditory compared to visual intervals could be established for extremely brief and longer intervals. However, when performance on duration discrimination of longer intervals

was controlled for modality-specific input from the sensory-automatic timing mechanism, the visual-auditory difference disappeared completely as indicated by virtually identical Weber fractions for both sensory modalities.

This pattern of results is consistent with the general notion that a 'hard' boundary between the sensory-automatic and the cognitive mechanism is rather unlikely to exist (cf. Rammsayer and Troche, 2014). Instead, it is reasonable to assume a transition zone from one timing mechanism to the other with a significant degree of processing overlap (Hellström and Rammsayer, 2002; Buonomano et al., 2009; Rammsayer and Ulrich, 2011). With increasing interval duration, the transition from a modality-specific, sensory-automatic to a more cognitive, amodal timing mechanism gets started. Within this transition zone, both mechanisms operate simultaneously but the influence of the sensory-automatic timing mechanism is expected to decrease with increasing interval duration. This decreasing influence of the sensory-automatic timing mechanism can account for the visual-auditory difference becoming gradually smaller with increasing interval duration. Converging evidence for this notion comes from Rammsayer and Ulrich's (2012) study where the visual-auditory difference was examined for standard durations ranging from 50 to 1400 ms. In this study, for brief standard durations below 800 ms, the visual-auditory difference, as indicated by Weber fractions, increased from 0.06 to 0.37 with standard durations decreasing from 800 to 50 ms. On the other hand, for standard durations longer than 800 ms, visual-auditory differences in Weber fractions remained almost constant at about 0.06. This gradient of visual-auditory differences in Weber fractions as a function of standard duration may be indicative of a transition from a purely modality-specific, sensory-automatic to a more cognitive, amodal timing mechanism. Moreover, these marked changes in visual-auditory differences as a function of interval duration observed in the present study and, in particular, those reported by Rammsayer and Ulrich (2012) clearly argue against the notion of a single, unitary timing mechanism as proposed by the Common Timing Hypothesis. Also neurophysiological data provided additional evidence in favor of both modality-specific and amodal mechanisms underlying the timing of intervals in the subsecond and second range (for concise reviews see Bueti, 2011; Wiener et al., 2011).

To date, the mechanism underlying the observed visual-auditory difference in duration discrimination of extremely brief intervals in the 10s-of-ms range still remains unclear. One notion refers to a finer temporal resolution due to more neural pulses accumulated with auditory intervals than with visual ones (e.g., Wearden et al., 1998; Penney et al., 2000; Droit-Volet et al., 2004). It is difficult to imagine, however, that the clock-like internal timing mechanism ticks so much faster for auditory than for visual intervals to completely account for a lowering in Weber fraction from 0.60 for visual to 0.17 for auditory intervals, as observed in the present study. This much higher temporal sensitivity in the auditory compared to the visual modality at the level of sensory-automatic temporal processing could also be due to less neural noise and, thus, faster and more accurate processing of auditory as compared to visual information (for a concise review see Stauffer et al., 2012).

A large number of studies on interval timing applying a dual-task approach support the view that processing of temporal information in the range of seconds occurs in working memory (e.g., Rammsayer and Lima, 1991; Fortin et al., 1993; Zakay, 1993; Sawyer et al., 1994; Fortin and Breton, 1995; Brown, 1997; Fortin, 1999; Rammsayer and Ulrich, 2011). Within the framework of the classical working memory model (e.g., Baddeley and Hitch, 1974; Baddeley, 1992, 2010), auditory and visual stimuli are assumed to be represented in separate and independent modality-specific stores. Quite obviously, this notion is at variance with the idea of an *amodal*, cognitive mechanism for temporal processing of longer intervals. In a most recent series of experiments, however, Salmela et al. (2014) provided experimental evidence that working memory resources are shared across representations in the auditory and visual sensory modalities. Thus, working memory can be considered a domain-general resource pool that is shared across modalities which is consistent with the basic assumption of an amodal, cognitive representation of time at a higher level of information processing (Stauffer et al., 2012; Filippopoulos et al., 2013).

Taken together, our findings are consistent with the general notion of two dissociable timing mechanisms underlying the obtained pattern of visual-auditory differences in duration

discrimination of extremely brief intervals in the 10s-of-s range and longer intervals in the 1-s range: a modality-specific, sensory-automatic and an amodal, cognitive mechanism. Most importantly, however, the marked visual-auditory differences observed for duration discrimination of extremely brief intervals appeared to depend on the predominating sensory-automatic temporal processing system. Only with increasing interval duration, the amodal, cognitive timing mechanism progressively contributes to the timing process. The present study also showed that it is possible to dissociate the contribution of the sensory-automatic timing system from that of the amodal, cognitive timing system. Finally, unlike the Distinct Timing Hypothesis in its strict sense, our findings argue for a transition zone characterized by a sensory-automatic and cognitive processing overlap. From this perspective, temporal processing of longer intervals in the 1-s range seems to be controlled by and functionally related to both sensory-automatic and cognitive timing mechanisms. As the evidence that the amodal, cognitive mechanism is impacted by the modality-specific, sensory-automatic timing mechanism is based on a null result, future studies are needed to provide additional converging evidence for this notion.

## REFERENCES

- Allman, M. J., Teki, S., Griffiths, T. D., and Meck, W. H. (2014). Properties of the internal clock: first- and second-order principles of subjective time. *Annu. Rev. Psychol.* 65, 743–771. doi: 10.1146/annurev-psych-010213-115117
- Baddeley, A. D. (1992). *Working Memory*. New York, NY: Oxford University Press.
- Baddeley, A. D. (2010). Working memory. *Curr. Biol.* 20, R136–R140. doi: 10.1016/j.cub.2009.12.014
- Baddeley, A. D., and Hitch, G. (1974). "Working memory," in *The Psychology of Learning and Motivation*, ed. G. H. Bower (New York, NY: Academic Press), 47–89.
- Brown, S. W. (1997). Attention resources in timing: interference effects in concurrent temporal and nontemporal working memory tasks. *Percept. Psychophys.* 59, 1118–1140. doi: 10.3758/BF03205526
- Bueti, D. (2011). The sensory representation of time. *Front. Integr. Neurosci.* 5:34. doi: 10.3389/fnint.2011.00034
- Buonomano, D. V., Bramen, J., and Khodadadifar, M. (2009). Influence of the interstimulus interval on temporal processing and learning: testing the state-dependent network model. *Philos. Trans. R. Soc. B* 364, 1865–1873. doi: 10.1098/rstb.2009.0019
- Church, R. M. (2003). "A concise introduction to scalar timing theory," in *Functional and Neural Mechanisms of Interval Timing*, ed. W. H. Meck (Boca Raton, FL: CRC Press), 3–32.
- Creelman, C. D. (1962). Human discrimination of auditory duration. *J. Acoust. Soc. Am.* 34, 582–593. doi: 10.1121/1.1918172
- Dixon, W. J. (1988). *BMDP Statistical Software Manual*, Vol. 1. Los Angeles, CA: University of California Press.
- Droit-Volet, S., Tourret, S., and Wearden, J. (2004). Perception of the duration of auditory and visual stimuli in children and adults. *Q. J. Exp. Psychol.* 57A, 797–818. doi: 10.1080/02724980343000495
- Filippopoulos, P. C., Hallworth, P., Lee, S., and Wearden, J. H. (2013). Interference between auditory and visual duration judgements suggests a common code time. *Psychol. Res.* 77, 708–715. doi: 10.1007/s00426-012-0464-6
- Fortin, C. (1999). "Short-term memory in temporal interval production," in *Short-Term/Working Memory: A Special Issue of the International Journal of Psychology*, eds I. Neath, G. D. A. Brown, M. Poirier, and C. Fortin (Hove: Psychology Press).
- Fortin, C., and Breton, R. (1995). Temporal interval production and processing in working memory. *Percept. Psychophys.* 57, 203–215. doi: 10.3758/BF03206507
- Fortin, C., Rousseau, R., Bourque, P., and Kirouac, E. (1993). Time estimation and concurrent nontemporal processing: specific interference from short-term-memory demands. *Percept. Psychophys.* 53, 536–548. doi: 10.3758/BF03211746
- Gibbon, J., and Church, R. M. (1984). "Sources of variance in an information processing theory of timing," in *Animal Cognition*, eds H. L. Roitblat, T. G. Bever, and H. S. Terrace (Hillsdale, NJ: Lawrence Erlbaum), 465–488.
- Gibbon, J., Malapani, C., Dale, C. L., and Gallistel, C. (1997). Toward a neurobiology of temporal cognition: advances and challenges. *Curr. Opin. Neurobiol.* 7, 170–184. doi: 10.1016/S0959-4388(97)80005-0
- Goldstone, S., and Lhamon, W. T. (1974). Studies of auditory-visual differences in duration judgment: 1. Sounds are judged longer than lights. *Percept. Mot. Skills* 39, 63–82. doi: 10.2466/pms.1974.39.1.63
- Grondin, S. (2003). "Sensory modalities and temporal processing," in *Time and Mind II: Information Processing Perspectives*, ed. H. Helfrich (Göttingen: Hogrefe & Huber Publishers), 61–77.
- Grondin, S. (2010). Timing and time perception: a review of recent behavioral and neuroscience findings and theoretical directions. *Atten. Percept. Psychophys.* 72, 561–582. doi: 10.3758/APP.72.3.561
- Grondin, S. (2014). About the (non)scalar property for time perception. *Adv. Exp. Med. Biol.* 829, 17–32. doi: 10.1007/978-1-4939-1782-2\_2
- Grondin, S., Meilleur-Wells, G., and Lachance, R. (1999). When to start explicit counting in time-intervals discrimination task: a critical point in the timing process of humans. *J. Exp. Psychol. Hum. Percept. Perform.* 25, 993–1004.
- Grondin, S., Ouellet, B., and Roussel, M.-E. (2004). Benefits and limits of explicit counting for discriminating temporal intervals. *Can. J. Exp. Psychol.* 58, 1–12. doi: 10.1037/h0087436
- Hellström, Å., and Rammsayer, T. (2002). "Mechanisms behind discrimination of short and long auditory durations," in *Annual Meeting of the International Society for Psychophysics*, eds J. D. da Silva, E. H. Matsuhashima, and N. P. Ribeiro-Filho (Rio de Janeiro: The International Society for Psychophysics), 110–115.
- Kaernbach, C. (1991). Simple adaptive testing with the weighted up-down method. *Percept. Psychophys.* 49, 227–229. doi: 10.3758/BF03214307
- Killeen, P. R., and Weiss, N. A. (1987). Optimal timing and the Weber function. *Psychol. Rev.* 94, 455–468. doi: 10.1037/0033-295X.94.4.455

- Kirk, R. E. (1995). *Experimental Design: Procedures for the Behavioral Sciences*. Pacific Grove, CA: Brooks/Cole.
- Lee, W. H. (1975). *Experimental Design and Analysis*. San Francisco, CA: Freeman.
- Lewis, P. A., and Miall, R. C. (2003). Distinct systems for automatic and cognitively controlled time measurement: evidence from neuroimaging. *Curr. Opin. Neurobiol.* 13, 250–255. doi: 10.1016/S0959-4388(03)00036-9
- Lewis, P. A., and Miall, R. C. (2006). A right hemispheric prefrontal system for cognitive time measurement. *Behav. Process.* 71, 226–234. doi: 10.1016/j.beproc.2005.12.009
- Lewis, P. A., and Miall, R. C. (2009). The precision of temporal judgement: milliseconds, many minutes, and beyond. *Philos. Trans. R. Soc. B* 364, 1897–1905. doi: 10.1098/rstb.2009.0020
- Luce, R. D., and Galanter, E. (1963). “Discrimination,” in *Handbook of Mathematical Psychology*, eds R. D. Luce, R. R. Bush, and E. Galanter (New York, NY: Wiley), 191–243.
- Merchant, H., Grahn, J., Trainor, L., Rohrmeier, M., and Fitch, W. T. (2015). Finding the beat: a neural perspective across humans and non-human primates. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 370, 20140093. doi: 10.1098/rstb.2014.0093
- Merchant, H., Harrington, D. L., and Meck, W. H. (2013a). Neural basis of the perception and estimation of time. *Annu. Rev. Neurosci.* 36, 313–336. doi: 10.1146/annurev-neuro-062012-170349
- Merchant, H., Pérez, O., Zarco, W., and Gámez, J. (2013b). Interval tuning in the primate medial premotor cortex as a general timing mechanism. *J. Neurosci.* 33, 9082–9096. doi: 10.1523/JNEUROSCI.5513-12.2013
- Michon, J. A. (1985). “The compleat time experimenter,” in *Time, Mind, and Behavior*, eds J. A. Michon and J. L. Jackson (Berlin: Springer), 21–52.
- Münsterberg, H. (1889). *Beiträge zur Experimentellen Psychologie: Heft 2*. Freiburg: Akademische Verlagsbuchhandlung von J. C. B. Mohr.
- Ortega, L., Lopez, F., and Church, R. M. (2009). Modality and intermittency effects on time estimation. *Behav. Process.* 81, 270–273. doi: 10.1016/j.beproc.2009.02.009
- Penney, T. B. (2003). “Modality differences in interval timing: attention, clock speed, and memory,” in *Functional and Neural Mechanisms of Interval Timing*, ed. W. H. Meck (Boca Raton, FL: CRC Press), 209–233.
- Penney, T. B., Gibbon, J., and Meck, W. H. (2000). Differential effects of auditory and visual signals on clock speed and temporal memory. *J. Exp. Psychol. Hum. Percept. Perform.* 26, 1770–1787.
- Penney, T. B., and Tournet, S. (2005). Les effets de la modalité sensorielle sur la perception du temps. *Psychol. Française* 50, 131–143. doi: 10.1016/j.psfr.2004.10.011
- Rammsayer, T. H. (1996). “Experimental evidence for different timing mechanisms underlying temporal discrimination,” in *Proceedings of the Twelfth Annual Meeting of the International Society for Psychophysics: Fechner Day 96*, ed. S. C. Masin (Padua: The International Society for Psychophysics), 63–68.
- Rammsayer, T. H. (2008). “Neuropharmacological approaches to human timing,” in *Psychology of Time*, ed. S. Grondin (Bingley: Emerald), 295–320.
- Rammsayer, T. H. (2012). Developing a psychophysical measure to assess duration discrimination in the range of milliseconds: methodological and psychometric issues. *Eur. J. Psychol. Assess.* 28, 172–180. doi: 10.1027/1015-5759/a000124
- Rammsayer, T. H. (2014). The effects of type of interval, sensory modality, base duration, and psychophysical task on the discrimination of brief time intervals. *Atten. Percept. Psychophys.* 76, 1185–1196. doi: 10.3758/s13414-014-0655-x
- Rammsayer, T. H., and Grondin, S. (2000). “Psychophysics of human timing,” in *Time and the Brain*, ed. R. Miller (Reading: Harwood Academic Publishers), 157–167.
- Rammsayer, T. H., and Lima, S. D. (1991). Duration discrimination of filled and empty auditory intervals: cognitive and perceptual factors. *Percept. Psychophys.* 50, 565–574. doi: 10.3758/BF03207541
- Rammsayer, T. H., and Troche, S. J. (2014). In search of the internal structure of the processes underlying interval timing in the sub-second and second range: a confirmatory factor analysis approach. *Acta Psychol.* 147, 68–74. doi: 10.1016/j.actpsy.2013.05.004
- Rammsayer, T. H., and Ulrich, R. (2001). Counting models of temporal discrimination. *Psychon. Bull. Rev.* 8, 270–277. doi: 10.3758/BF03196161
- Rammsayer, T. H., and Ulrich, R. (2011). Elaborative rehearsal of nontemporal information interferes with temporal processing of durations in the range of seconds but not milliseconds. *Acta Psychol.* 137, 127–133. doi: 10.1016/j.actpsy.2011.03.010
- Rammsayer, T. H., and Ulrich, R. (2012). The greater temporal acuity in the reminder task than in the 2AFC task is independent of standard duration and sensory modality. *Can. J. Exp. Psychol.* 66, 26–31. doi: 10.1037/a0025349
- Salmela, V. R., Moisaala, M., and Alho, K. (2014). Working memory resources are shared across sensory modalities. *Atten. Percept. Psychophys.* 76, 1962–1974. doi: 10.3758/s13414-014-0714-3
- Sawyer, T. F., Meyers, P. J., and Huser, S. J. (1994). Contrasting task demands alter the perceived duration of brief time intervals. *Percept. Psychophys.* 56, 649–657. doi: 10.3758/BF03208358
- Spencer, R. M. C., Karmarkar, U., and Ivry, R. B. (2009). Evaluating dedicated and intrinsic models of temporal encoding by varying context. *Philos. Trans. R. Soc. B* 364, 1853–1863. doi: 10.1098/rstb.2009.0024
- Stauffer, C., Haldemann, J., Troche, S. J., and Rammsayer, T. H. (2012). Auditory and visual temporal sensitivity: evidence for a hierarchical structure of modality-specific and modality-independent levels of temporal information processing. *Psychol. Res.* 76, 20–31. doi: 10.1007/s00426-011-0333-8
- Tabachnick, B. G., and Fidell, L. S. (2013). *Using Multivariate Statistics*. Upper Saddle River, NJ: Pearson.
- Treisman, M. (1963). Temporal discrimination and the indifference interval: implications for a model of the “internal clock.” *Psychol. Monogr.* 77, 1–31. doi: 10.1037/h0093864
- van Wassenhove, V. (2009). Minding time in an amodal representational space. *Philos. Trans. R. Soc. B* 364, 1815–1830. doi: 10.1098/rstb.2009.0023
- Walker, J. T., and Scott, K. J. (1981). Auditory-visual conflicts in the perceived duration of lights, tones, and gaps. *J. Exp. Psychol. Hum. Percept. Perform.* 7, 1327–1339.
- Wearden, J. H., Edwards, H., Fakhri, M., and Percival, A. (1998). Why “sounds are judged longer than lights”: application of a model of the internal clock in humans. *Q. J. Exp. Psychol. B Comp. Physiol. Psychol.* 51, 97–120.
- Wiener, M., Matell, M. S., and Coslett, H. B. (2011). Multiple mechanisms for temporal processing. *Front. Psychol.* 5:31. doi: 10.3389/fnint.2011.00031
- Zakay, D. (1993). The roles of nontemporal information processing load and temporal expectations in children’s prospective time estimation. *Acta Psychol.* 84, 271–280. doi: 10.1016/0001-6918(93)90064-X

**Conflict of Interest Statement:** 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.

Copyright © 2015 Rammsayer, Bortner and Troche. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





# Odors Bias Time Perception in Visual and Auditory Modalities

Zhenzhu Yue<sup>1\*</sup>, Tianyu Gao<sup>1</sup>, Lihan Chen<sup>2,3</sup> and Jiashuang Wu<sup>1</sup>

<sup>1</sup> Department of Psychology, Sun Yat-sen University, Guangzhou, China, <sup>2</sup> Department of Psychology and Beijing Key Laboratory of Behavior and Mental Health, Peking University, Peking, China, <sup>3</sup> Key Laboratory of Machine Perception (Ministry of Education), Peking University, Beijing, China

## OPEN ACCESS

### Edited by:

Marc Wittmann,  
Institute for Frontier Areas of  
Psychology and Mental Health,  
Germany

### Reviewed by:

Devin Terhune,  
University of Oxford, UK  
Lukasz Smigielski,  
University of Zurich, Switzerland

### \*Correspondence:

Zhenzhu Yue  
yuezhzhu@gmail.com

### Specialty section:

This article was submitted to  
Perception Science,  
a section of the journal  
Frontiers in Psychology

**Received:** 04 July 2015

**Accepted:** 31 March 2016

**Published:** 22 April 2016

### Citation:

Yue Z, Gao T, Chen L and Wu J (2016)  
Odors Bias Time Perception in Visual  
and Auditory Modalities.  
Front. Psychol. 7:535.  
doi: 10.3389/fpsyg.2016.00535

Previous studies have shown that emotional states alter our perception of time. However, attention, which is modulated by a number of factors, such as emotional events, also influences time perception. To exclude potential attentional effects associated with emotional events, various types of odors (inducing different levels of emotional arousal) were used to explore whether olfactory events modulated time perception differently in visual and auditory modalities. Participants were shown either a visual dot or heard a continuous tone for 1000 or 4000 ms while they were exposed to odors of jasmine, lavender, or garlic. Participants then reproduced the temporal durations of the preceding visual or auditory stimuli by pressing the spacebar twice. Their reproduced durations were compared to those in the control condition (without odor). The results showed that participants produced significantly longer time intervals in the lavender condition than in the jasmine or garlic conditions. The overall influence of odor on time perception was equivalent for both visual and auditory modalities. The analysis of the interaction effect showed that participants produced longer durations than the actual duration in the short interval condition, but they produced shorter durations in the long interval condition. The effect sizes were larger for the auditory modality than those for the visual modality. Moreover, by comparing performance across the initial and the final blocks of the experiment, we found odor adaptation effects were mainly manifested as longer reproductions for the short time interval later in the adaptation phase, and there was a larger effect size in the auditory modality. In summary, the present results indicate that odors imposed differential impacts on reproduced time durations, and they were constrained by different sensory modalities, valence of the emotional events, and target durations. Biases in time perception could be accounted for by a framework of attentional deployment between the inducers (odors) and emotionally neutral stimuli (visual dots and sound beeps).

**Keywords:** time perception, odor, visual, auditory, adaptation

## INTRODUCTION

Time perception is an important aspect of human life. Although time perception is an important ability for human survival, people often overestimate or underestimate the actual duration of events. Both temporal and non-temporal factors contribute to biases in time estimation. One famous example of temporal factors is from *Vierordt's Law*, which states that judgments of relatively short time intervals are lengthened while the judgments of relatively long time intervals are

shortened (Buetti et al., 2008; Block and Gruber, 2014). Non-temporal factors that affect time perception include sensory modality (Goldstone and Lhamon, 1974; Gruber and Block, 2013), emotion states (Droit-Volet and Meck, 2007; Noulhiane et al., 2007; Tipples, 2008; Wittmann and Paulus, 2008; Droit-Volet and Gil, 2009; Gil and Droit-Volet, 2011; Lee et al., 2011), dynamic features of stimuli (Kanai et al., 2006), and directions of motion stimuli (Ono and Kitazawa, 2010). For the modality effect, previous studies have shown that individuals tend to perceive durations as longer in the auditory modality than in the visual modality when the physical durations are less than 1 s (Goldstone and Lhamon, 1974; Wearden et al., 1998). However, the difference between time estimations in the visual and auditory modalities decreased or disappeared when the stimulus duration was longer than 3–5 s (Gruber and Block, 2013; Block and Gruber, 2014). Those differential effects indicate that the illusory bias in time perception is duration-selective.

Among the non-temporal factors, attentional factors and their modulations of the “internal clock” have mostly been exploited to account for the bias in time perception. The traditional internal clock model, which mainly consists of a pacemaker and an accumulator as well as devices of memory and decision components, could explain the processing of time estimation (Gibbon et al., 1984; Buhusi and Meck, 2005). The pacemaker sends out pulses (that is, units of elapsed time) to the accumulator at a particular rate, and the subjectively perceived duration of time is defined by the number of temporal units accumulated over an actual time interval (Schwartz et al., 1986). A close inspection of the internal-clock model suggests that attention and arousal states modulate the accumulation of pulses and cause variations in subjective time estimation (Wittmann and Paulus, 2008). On one hand, increased attentional focus on time perception led to an accumulation of more pulses (Schreuder et al., 2014). On the contrary, when attention was attracted by other task-irrelevant factors, fewer pulses were calculated, and the perceived duration was thus shorter for a given time interval (Droit-Volet and Meck, 2007). On the other hand, increased arousal led to the increased rate of pulses emitted by the pacemaker and induced a greater/faster accumulation of pulses over time. Thus, a given time interval tended to be perceived as longer in a high than in a low arousal condition.

By using emotional faces, Zhang and Zhou (2007) investigated the influence of emotional events on time perception. They found that participants underestimated the duration of angry faces but overestimated the duration of happy faces. Although, arousal states were used to explain their results, their results could also be explained by the distribution of attentional resources between temporal and non-temporal processing. Processing emotional information and estimating time intervals share common attentional resources. When emotional events captured attention, attentional resources allocated for processing time information were reduced. Hence, the perceived subjective time was shorter than the actual duration due to the loss of pacemaker pulses. In the above paradigm, the target stimuli for time estimation coupled emotional information and attentional factors, which made it difficult to tease apart the roles of the different variables (attention vs. emotion) and to exclude the potential confounding

variables (such as “emotional states”) induced by the target stimuli themselves.

To overcome this potential confound, an ideal experimental design would be to implement/modulate the arousal states from a third sensory modality while investigating time perception in the target modality. To achieve this, in the current study, we investigated time perception in visual and auditory modalities while manipulating the arousal states in a third modality, namely olfaction, to minimize the emergent properties of emotional information associated with the targets and to examine the other modulating factors beyond arousal states that would affect time perception.

It has been well documented that odors can induce different arousal experiences. For example, odors with positive emotional experience, such as lavender, chamomile, and sandalwood, can decrease anxiety levels (Schwartz et al., 1986; Roberts and Williams, 1992; Moss et al., 2003). In contrast, jasmine and rosemary have been shown to improve alertness and enhance cognitive performance (Kovar et al., 1987; Diego et al., 1998). By using a priming paradigm, Gros et al. (2015) investigated the influence of emotional prime stimuli on the duration estimation of a target. Participants estimated the duration of a pure sound, which was primed by odors or emotional videos. Their results showed that odors consistently activated the arousal system because the measured skin conductance (SC) increased consistently, and no decrease in SC was observed across time. Their results suggest that odors could be well used to investigate the arousal-related mechanism. However, a previous study has explored the effect of odor on time perception (Schreuder et al., 2014), and a time distortion was still observed even though no increase in arousal was indicated by SC or heart rate. In this study, participants were assigned to the rosemary (arousing), peppermint (relaxing), or no odor (control) condition, and they sat either upright (arousing) or lied down (relaxing) during the time perception task. Participants estimated the lengths of time intervals (1.33, 1.58, and 2.17 min) and produced the durations by clicking a mouse button twice to mark the beginning and end of the time periods. Their results showed that the participants produced shorter time intervals in the rosemary odor condition than in the no odor condition, suggesting that odors impacted time perception. However, it should be noted that all time intervals used in this experiment exceeded 1 min, which made the exploration of the timing mechanism illusive when the target time interval was shorter. According to Fraisse (1984), estimating time duration longer than 5 s would mainly exploit a cognitive mechanism and need long-term memory. Therefore, it is reasonable to assume that estimating time durations of less than 5 s might tap into different cognitive resources/processes and hence have different behavioral patterns than comparing time estimations of long durations, as was conducted in Schreuder and colleagues’ study (2014). Therefore, we aimed to investigate how odors influenced estimates of time duration of less than 5 s (Poeppel, 2004) and explored different timing mechanisms within 5 s.

For time perception, although some studies have found that there are differences between auditory and visual signals (Goldstone and Lhamon, 1974; Wearden et al., 1998), others

have not found modality differences (Bobko et al., 1977). For example, Penney et al. (2000) investigated the effect of stimulus modality on duration classification with a duration bisection task (Allan and Gibbon, 1991). Visual or auditory signals were timed either simultaneously on some trials or alone on other trials. In the training period, participants were presented either short or long anchor durations of signals. Participants made duration judgments (short or long) in the testing period in which there were two anchors and five geometrically spaced intermediate probe durations. They found that modality effect was only observed in blocks containing only a single modality condition, but it was not observed when participants experienced both modalities in the same block. Their results indicated that the temporal precision across sensory modalities is different (Welch and Warren, 1980), and an internal clock runs at a faster rate for auditory than for visual signals. The main purpose of the present study was to investigate to what extent the perception of visual or auditory stimulus duration was influenced by the presence of odors within the time range less than 5 s. To achieve this purpose, two odors of positive affect (jasmine- high arousal; lavender- low arousal) and one odor of negative affect (garlic- high arousal) were used. To increase the accuracy of duration reproductions, a sample time interval was presented first, and then participants produced the same time intervals in the present study. Participants were shown a dot or heard a tone for either 1000 or 4000 ms. After the stimulus, participants estimated the stimulus duration by pressing the space bar twice to demarcate the beginning and end of a produced time duration. We hypothesized that the arousal level induced by the odors would affect the perceived duration in both the visual and auditory modalities. Individuals perceived a given time interval as longer in the high arousal condition than in the low arousal condition (Tremblay and Fortin, 2003) because the arousal states quicken the accumulation of pulses. Moreover, the perceived duration was also influenced by the attention mechanism when attentional resources were not depleted and could be directed to the targets because olfactory stimuli were presented simultaneously.

As in the other sensory modalities, exposure to the odors for a long time period would lead to sensory adaptation and change the subjective sensitivities to the odors. If any emotional states were triggered by the odors, they would influence the time perception for target events as a function of the passage of time. Hence, we compared performances of the initial and the final parts in the experiment, which were separated from each other by approximately 7–10 min, to show the potential bias in time across the different adaptation phases.

## MATERIALS AND METHODS

### Participants

One hundred undergraduate students (29 males, 71 females) from Sun Yat-sen University participated in this study. They were 17–24 years old (Mean age = 19.7, SD = 1.40). Data from two participants were excluded due to exceeding three standard deviations of the average. Therefore, the final analysis consisted of data from 98 participants, including 23 participants (10 males) in the jasmine condition, 23 participants (8 males) in the lavender

condition, 24 participants (6 males) in the garlic condition, and 28 participants (5 males) in the no odor condition.

All participants were right handed except for one. Participants self-reported normal olfaction, audition and normal or corrected-to-normal vision. They were paid 10 Chinese yuan for taking part in the study. The study was conducted in accordance with the guidelines in the Declaration of Helsinki (2000) and was approved by the Ethics Committee of Department of Psychology, Sun Yat-sen University. All participants gave their written informed consent before taking part in the study.

### Stimuli

#### Olfactory Stimuli

Three odors (garlic, jasmine, and lavender) were used. The garlic odor was picked from a solution prepared by dissolving 525 g of garlic odorizer powder into 300 ml of water. The jasmine and lavender odors were made from 300 ml of liquid air fresheners with the respective fragrance. No negative low arousal odor was used because we were focusing on the categories of “positive” and “negative” odors. Moreover, most negative odors are highly arousing, and it is not convenient to modulate the level of arousal state with negative odors.

To avoid the mixing arousal induced by different odors, only one of the three odors was randomly assigned to each participant. We soaked two cotton pads ( $6 \times 5 \text{ cm}^2$ ) in 5 ml of an odor liquid for 20 min. Before the experiment, the experimenter brought the cotton pads into the room and fixed them under the desk. We occluded the cotton pads such that participants only smelled them but could not see them. After the cotton pads had been placed in the room for 30 min, participants entered the room ( $1.2 \times 1.7 \text{ m}^2$ ) to start the experiment.

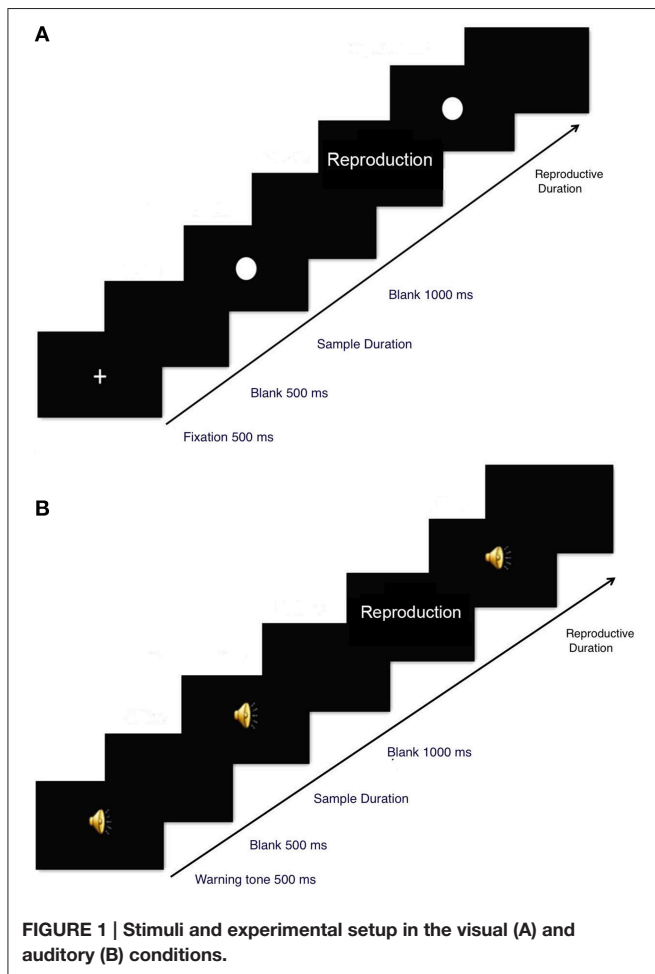
#### Visual and Auditory Stimuli

The visual stimulus was a white dot (of visual angle  $5.27^\circ \times 5.27^\circ$ , Luminance  $10.4 \text{ cd/m}^2$ ), which was presented on a 17-inch monitor (Refresh rate 60 Hz) and controlled by E-prime (<http://www.pstnet.com/eprime.cfm>). The auditory stimulus was a pure tone (500 Hz, 70 dB) presented via headphones (EDIFIER H850) to both ears.

### Procedures

Participants sat in front of a monitor in a room, which was dimly lit and windowless. The viewing distance was 60 cm. Participants did not receive any information about the odors before the experiment.

In the visual condition, a fixation cross (of visual angle  $5.27^\circ \times 5.27^\circ$ ) was presented in the center of the monitor for 500 ms followed by a 500 ms blank (see **Figure 1A**). Next, a white dot was presented in the center of the screen for an average of 1000 ms (randomly selected from 800, 900, 1000, 1100, or 1200 ms) or for an average of 4000 ms (randomly selected from 3800, 3900, 4000, 4100, or 4200 ms). Each duration was presented six times in each condition. Then, the screen turned black and the participant waited for 1000 ms before (s)he reproduced the presentation duration of the white dot. Participants could reproduce the time interval after the word “reproduction” was present on the screen, and the word was kept on the screen until the produced duration



was finished. Specifically, when making a response, a participant first pressed the spacebar once, and then a white dot appeared on the screen. (S)he waited for an equivalent length of time that (s)he believed the original visual stimulus duration to be and then pressed the spacebar again to end the trial. The screen then turned black for 1000 ms before a new trial began.

In the auditory condition, a cueing sound (2000 Hz, 500 ms) was delivered via a set of headphones, which was followed by a 500 ms blank (see **Figure 1B**). Next, a pure tone (500 Hz, 70 dB) was presented for an average of 1000 ms (randomly selected from 800, 900, 1000, 1100, or 1200 ms) or 4000 ms (randomly selected from 3800, 3900, 4000, 4100, or 4200 ms). Participants then waited quietly for 1000 ms and reproduced the duration of the auditory stimulus by pressing the spacebar twice (as in the procedure in visual condition). When participants first pressed the spacebar, a pure tone initiated and lasted until participants pressed the space bar again.

Each participant completed two blocks with visual stimuli and two blocks with auditory stimuli, each consisting of 30 trials. Half of the participants started with the visual task, and the other half started with the auditory task (block orders were counterbalanced between subjects). For each condition, at least 10 practice trials

were completed prior to the start of the formal experiment. Participants took a break after the completion of each block.

After the time reproduction task, participants answered the following survey of four questions in the same room: (1) Did you notice the odor in the room? (2) Please specify the type of odor in the room: jasmine, lavender, or garlic. (3) When you smell this odor, please rate your emotional experience on a scale from  $-4$ (extremely unpleasant) to  $4$ (extremely pleasant). (4) When you smell this odor, please rate your emotional experience on a scale from  $-4$ (extremely calm) to  $4$ (extremely aroused). After a participant answered the questions and left the room, we discarded the cotton pads and ventilated the room for 20 min.

## Data Analyses

For the odors used in the present study, participants rated each odor according to its valence and arousal. The olfactory discrimination and evaluation were analyzed with one-way analyses of variance (ANOVA).

For the time reproduction task, the dependent variable was the difference between the estimated duration and the actual duration. A positive value meant longer reproductions of the time interval than the actual duration, and a negative value meant shorter reproductions of the time interval than the actual duration. A ratio score was also calculated with the following formula:  $[T \text{ corrected score} = (T \text{ estimated} - T \text{ standard}) / T \text{ standard}]$  (Brown, 1985). We then performed a 4 (odor type: high-arousal positive odor: jasmine, low-arousal positive odor: lavender, high-arousal negative odor: garlic, and no odor)  $\times$  2 (modality: visual vs. auditory)  $\times$  2 (interval: short vs. long) ANOVA. The odor type was a between-subjects variable, while the other two were within-subjects variables.

To measure the effect of the adaption to odors, the performance in the time reproduction task was compared between the initial and the final block for each modality. A four-way repeated measures ANOVA was conducted, with another within-subjects variable, adaptation phase (initial block vs. final block), in addition to the above three factors: odor type, modality and interval.

## RESULTS

### Emotion Induction by Each Odor

Participants in the three odor conditions (not including those in the no-odor condition) rated the valence and arousal levels of each odor. All participants noticed the odor in the room and identified the odors correctly.

The valence and arousal scores were summarized in **Figure 2**. One-way ANOVA revealed a main effect of odor by valence,  $F_{(2, 67)} = 41.4, p < 0.0001, \eta^2 = 0.553$ , and a marginal main effect of odor on arousal,  $F_{(2, 67)} = 3.07, p = 0.053, \eta^2 = 0.084$ . For the emotional valence, further  $t$ -tests showed that the pleasantness of the jasmine odor ( $M = 1.30, SE = 0.34$ ) and lavender odor ( $M = 1.43, SE = 0.27$ ) were significantly greater than garlic odor ( $M = -1.66, SE = 0.21$ ),  $t_{(45)} = 7.53, p < 0.01$ , Cohen's  $d = 2.181$ , and  $t_{(45)} = 9.3, p < 0.01, d = 2.649$ , respectively. For the emotional arousal, the arousal of the garlic odor ( $M = 0.54, SE = 0.38$ ) was significantly greater than the arousal of the jasmine odor ( $M$



$= -0.65$ ,  $SE = 0.39$ ),  $t_{(45)} = 2.2$ ,  $p < 0.05$ ,  $d = 0.638$ , and the lavender odor ( $M = -0.61$ ,  $SE = 0.40$ ),  $t_{(45)} = 2.08$ ,  $p < 0.05$ ,  $d = 0.609$ . There was no difference between the jasmine and lavender odors for the emotional valence [ $t_{(44)} = -0.302$ ,  $p = 0.76$ ,  $d = 0.089$ ] or the emotional arousal [ $t_{(44)} = -0.078$ ,  $p = 0.94$ ,  $d = 0.021$ ].

## Time Reproduction Task

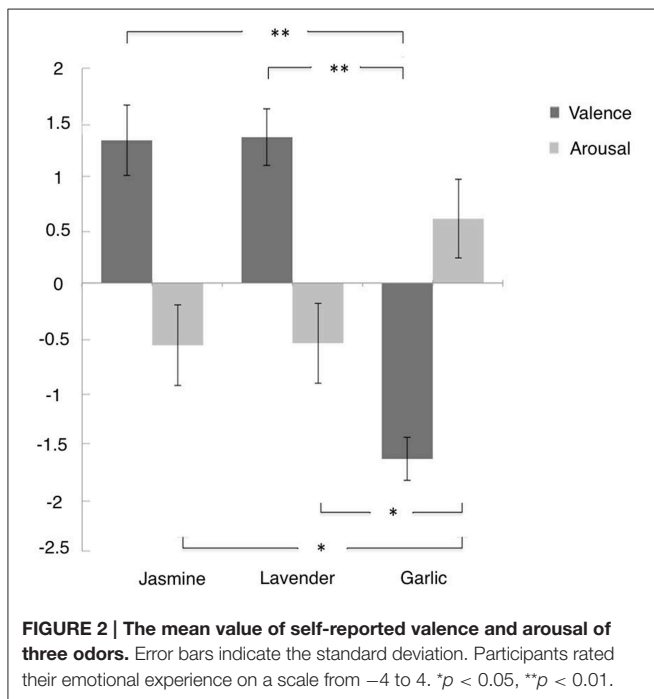
The outlier data of the participants, i.e., the reaction times exceeding three standard deviations (less than 5 %) in each experimental condition, were removed. **Table 1** shows the mean differences between the reproduced time intervals and actual time intervals in which the actual duration was subtracted from the reproduced duration. To control for the initial bias for the

baseline intervals (short vs. long), the following ratio score was also adopted:  $T \text{ corrected score} = (T \text{ estimated} - T \text{ standard})/T \text{ standard}$ .

A three-way ANOVA (odor  $\times$  interval  $\times$  modality) revealed a significant main effect of odor,  $F_{(3, 94)} = 2.92$ ,  $p < 0.05$ ,  $\eta^2 = 0.085$ . Further  $t$ -tests showed that the time estimation bias in the lavender condition ( $M = 92.2 \text{ ms}$ ,  $SE = 44.0$ ) was significantly greater than that in the jasmine condition ( $M = -71.36 \text{ ms}$ ,  $SE = 44.0$ ),  $t_{(44)} = -2.84$ ,  $p < 0.01$ ,  $d = 0.775$  and greater than that in the garlic condition ( $M = -58.5 \text{ ms}$ ,  $SE = 43.0$ ),  $t_{(45)} = 2.5$ ,  $p < 0.05$ ,  $d = 0.715$ . No significant differences between the no odor condition and the three odor conditions were found, and all  $p$ 's were greater than 0.8. In addition, the main effect of the interval was significant,  $F_{(1, 94)} = 120.75$ ,  $p < 0.01$ ,  $\eta^2 = 0.562$ . Participants tended to reproduce longer durations than the actual durations for short time intervals ( $M = 161.83 \text{ ms}$ ,  $SE = 21.76$ ) and reproduce shorter durations for long time intervals ( $M = -177.81 \text{ ms}$ ,  $SE = 30.32$ ) in all odor conditions, thus resembling *Vierordt's Law*. The main effect of modality did not reach significance,  $F_{(1, 94)} = 0.68$ ,  $p = 0.41$ .

The interaction between the interval and modality was significant,  $F_{(1, 94)} = 22.683$ ,  $p < 0.01$ ,  $\eta^2 = 0.194$  (see **Figure 3**). Further  $t$ -tests showed that for the short interval condition, the reproductions in the auditory modality ( $M = 207.23 \text{ ms}$ ,  $SE = 25.62$ ) were significantly longer than those in the visual modality ( $M = 117.08 \text{ ms}$ ,  $SE = 24.14$ ),  $t_{(97)} = 3.884$ ,  $p < 0.001$ ,  $d = 0.365$ . By contrast, for the long interval condition, the shorter reproductions of auditory time intervals ( $M = -207.63 \text{ ms}$ ,  $SE = 32.49$ ) was significantly greater than that of visual time intervals ( $M = -147.91 \text{ ms}$ ,  $SE = 35.21$ ),  $t_{(97)} = -2.069$ ,  $p < 0.05$ ,  $d = 0.178$ . The difference between the visual and auditory modalities ( $M = 90.15 \text{ ms}$ ,  $SE = 23.21$ ) for longer time intervals did not differ significantly from the difference between modalities for shorter intervals ( $M = 59.72 \text{ ms}$ ,  $SE = 28.87$ ),  $t_{(97)} = 0.738$ ,  $p > 0.05$ ,  $d = 0.117$ . The interaction between odor, interval and modality did not reach significance,  $F_{(3, 94)} = 2.15$ ,  $p = 0.09$ .

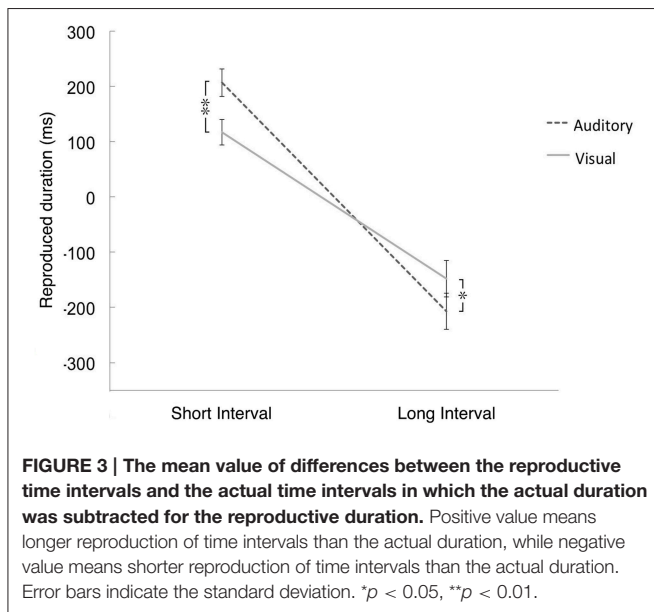
The analysis of variance (ANOVA) of the ratio scores showed a main effect of interval,  $F_{(1, 94)} = 90.832$ ,  $p < 0.001$ ,  $\eta^2 =$



**FIGURE 2 | The mean value of self-reported valence and arousal of three odors.** Error bars indicate the standard deviation. Participants rated their emotional experience on a scale from -4 to 4. \* $p < 0.05$ , \*\* $p < 0.01$ .

**TABLE 1 | The means (ms) and standard errors (se) of the over- or under-estimation of time intervals (the differences between reproductive time intervals and real time intervals) in all experimental conditions. The ratio scores were also calculated and shown in the table.**

Odor	Short interval		Long interval	
	Auditory	Visual	Auditory	Visual
<b>MEAN REPRODUCTION INTERVAL (ms)</b>				
Jasmine	183.7 (51.9)	35.5 (49.35)	-261.3 (67.33)	-243.3 (70.84)
Lavender	314.2 (51.9)	175.3 (49.35)	-125.1 (67.33)	-4.4 (70.84)
Garlic	134.7 (50.88)	96.4 (48.31)	-232.0 (65.91)	-233.1 (69.35)
No odor	200.8 (47.11)	154.0 (44.72)	-210.4 (61.02)	-121.6 (64.21)
<b>T CORRECTED SCORE = (MEAN REPRODUCTION INTERVAL - STANDARD INTERVAL)/STANDARD INTERVAL</b>				
Jasmine	0.185 (0.049)	0.025 (0.047)	-0.063 (0.016)	-0.054 (0.016)
Lavender	0.267 (0.049)	0.124 (0.047)	-0.029 (0.016)	-0.007 (0.016)
Garlic	0.18 (0.044)	0.143 (0.042)	-0.048 (0.014)	-0.018 (0.014)
No odor	0.108 (0.048)	0.059 (0.046)	-0.051 (0.015)	-0.046 (0.016)



0.491, and a significant main effect of modality,  $F_{(1, 94)} = 10.593$ ,  $p < 0.01$ ,  $\eta^2 = 0.101$ . The interaction between interval and modality was significant,  $F_{(1, 94)} = 25.045$ ,  $p < 0.001$ ,  $\eta^2 = 0.21$  (see Table 1, it has similar trend as in Figure 3). Further tests revealed that for the short interval condition, the ratio score for longer reproductions of the auditory time interval ( $M = 0.185$ ,  $SE = 0.024$ ) was significantly larger than that of the visual time interval ( $M = 0.088$ ,  $SE = 0.023$ ),  $t_{(97)} = 4.05$ ,  $p < 0.0001$ ,  $d = 0.417$ . By contrast, for the long interval condition, the ratio score for shorter reproductions of the auditory time interval ( $M = -0.048$ ,  $SE = 0.008$ ) was significantly larger than that of the visual time interval ( $M = -0.031$ ,  $SE = 0.008$ ),  $t_{(97)} = -2.727$ ,  $p < 0.01$ ,  $d = 0.215$ . Moreover, the difference between the visual and auditory modalities for longer time intervals ( $M = 0.097$ ,  $SE = 0.006$ ) was significantly larger than the difference between the two modalities for shorter intervals ( $M = 0.017$ ,  $SE = 0.023$ ),  $t_{(97)} = 4.831$ ,  $p < 0.001$ ,  $d = 0.169$ . However, the main effect of odor,  $F_{(3, 94)} = 1.965$ ,  $p = 0.125$ ,  $\eta^2 = 0.059$ , and the interaction between odor, interval and modality did not reach significance,  $F_{(3, 94)} = 1.795$ ,  $p = 0.153$ ,  $\eta^2 = 0.054$ .

## The Adaptation of Odors

The effect of odor adaptation on time perception was analyzed by comparing the performance of trials in the initial (the first block) and final (the fourth block) parts. The final experimental block started approximately 7 min after the end of the initial experimental block. The mean differences between the reproduced duration and the actual duration were calculated in each experimental condition. A four-way mixed ANOVA was conducted with the between-subjects factor of odor (jasmine, lavender, garlic, vs. no odor), the within-subjects factors of interval (short vs. long), modality (visual vs. auditory), and the adaptation phase (initial vs. final block). A significant main effect of odor was observed,  $F_{(3, 94)} = 3.341$ ,  $p < 0.001$ ,  $\eta^2 = 0.096$ , which was in accordance with our earlier results. The main effect of interval was significant,  $F_{(1, 94)} = 41.965$ ,  $p < 0.01$ ,  $\eta^2 = 0.309$ ,

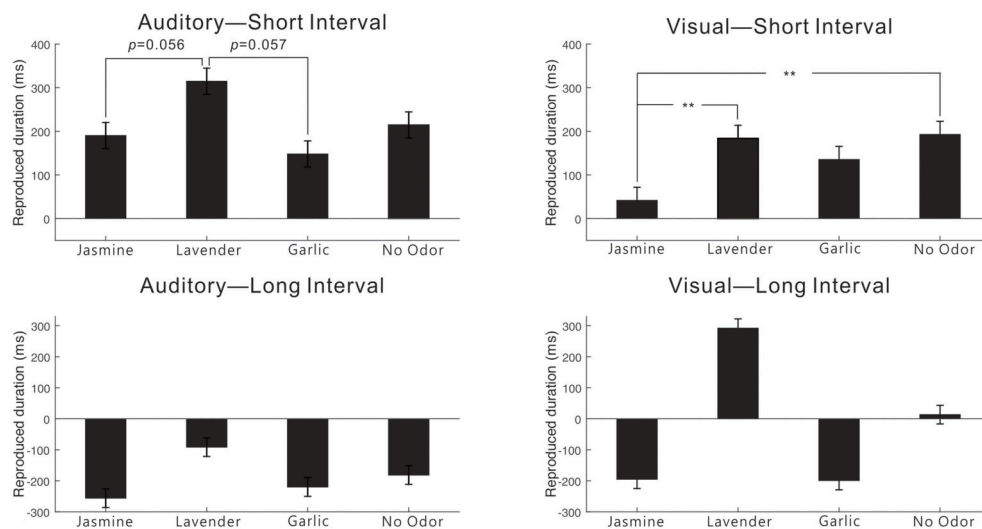
which suggested that participants produced longer durations than the actual durations for the short time intervals ( $M = 177.6$  ms,  $SE = 23.50$ ) and produced shorter durations for the long time intervals ( $M = -104.77$  ms,  $SE = 44.34$ ). The main effect of adaptation phase was significant,  $F_{(1, 94)} = 7.24$ ,  $p < 0.01$ ,  $\eta^2 = 0.072$ , which suggested that participants reproduced longer time durations than the actual duration in the final block ( $M = 84.2$  ms,  $SE = 40.58$ ) than in the initial block ( $M = -11.2$  ms,  $SE = 23.45$ ).

A significant interaction between modality and interval was observed,  $F_{(1, 94)} = 14.509$ ,  $p < 0.001$ ,  $\eta^2 = 0.134$  (see Figure 4). Further  $t$ -tests showed that in the short time interval condition, the magnitude of the longer reproduction of the auditory time interval ( $M = 216.87$  ms,  $SE = 26.53$ ) than the actual time interval was significantly larger than that of the visual time interval ( $M = 138.38$  ms,  $SE = 25.61$ ),  $t_{(97)} = 3.25$ ,  $p = 0.002$ ,  $d = 0.304$ . In contrast, in the long time interval trials, the magnitude of the shorter reproduction of the auditory time interval ( $M = -187.20$  ms,  $SE = 34.00$ ) than the actual time interval was significantly larger than that of the visual time interval ( $M = -22.12$  ms,  $SE = 69.88$ ),  $t_{(97)} = -2.532$ ,  $p < 0.05$ ,  $d = 0.303$ , which was consistent with our earlier results. There is a trend toward significance for the interaction between modality, adaptation phase and odor,  $F_{(3, 94)} = 2.282$ ,  $p = 0.084$ ,  $\eta^2 = 0.068$ . Further, analyses revealed a significant main effect of adaptation for the no odor condition only,  $F_{(1, 27)} = 4.361$ ,  $p < 0.05$ ,  $\eta^2 = 0.139$ , indicating that the longer reproduction of time intervals was larger in the final block ( $M = 110.162$  ms,  $SE = 74.34$ ) than in the initial block ( $M = 9.479$  ms,  $SE = 43.24$ ).

Importantly, there is a trend toward significance for the interaction between modality, interval and odor,  $F_{(3, 94)} = 2.339$ ,  $p = 0.078$ ,  $\eta^2 = 0.069$  (see Figure 4). For the short interval condition, further analyses showed a significant main effect of modality,  $F_{(1, 94)} = 12.043$ ,  $p < 0.01$ ,  $\eta^2 = 0.114$ , and a trend toward significance for the interaction between modality and odor,  $F_{(3, 94)} = 2.476$ ,  $p = 0.066$ ,  $\eta^2 = 0.073$ . Further,  $t$ -tests showed that for the auditory modality, the reproduction of the time interval in the lavender condition ( $M = 314.73$  ms,  $SE = 54.59$ ) was nearly significantly longer than that in the jasmine ( $M = 190.28$  ms,  $SE = 54.59$ ),  $t_{(44)} = -1.959$ ,  $p = 0.056$ ,  $d = 0.475$ , and garlic conditions ( $147.95$  ms,  $SE = 53.44$ ),  $t_{(45)} = 1.951$ ,  $p = 0.057$ ,  $d = 0.637$ . For the visual modality, the longer reproduction of the time interval than the actual duration in the jasmine condition ( $M = 41.66$  ms,  $SE = 52.69$ ) was significantly smaller than that in the lavender condition ( $M = 183.68$  ms,  $SE = 52.69$ ),  $t_{(44)} = -2.245$ ,  $p = 0.03$ ,  $d = 0.562$ , and no odor condition ( $M = 192.84$  ms,  $SE = 47.75$ ),  $t_{(49)} = -2.145$ ,  $p = 0.037$ ,  $d = 0.598$ . In contrast, for the long interval condition, only a significant main effect of odor,  $F_{(3, 94)} = 2.748$ ,  $p = 0.047$ ,  $\eta^2 = 0.081$ , and a significant main effect of modality,  $F_{(1, 94)} = 6.469$ ,  $p = 0.013$ ,  $\eta^2 = 0.064$ , were observed.

## DISCUSSION

The primary goal of the present study was to investigate the influence of different odors on time perception in both visual and auditory modalities. Moreover, we investigated whether the adaptation of odors affected perceived time of different ranges



**FIGURE 4 | The mean value of differences between the reproductive time intervals and the actual time intervals for both visual and auditory modalities during the experiment in which the actual duration was subtracted from the estimated duration. A positive value means a longer reproduction of time intervals than the actual duration, while a negative value means a shorter reproduction of time intervals than the actual duration. Error bars indicate the standard deviation. \*\* $p < 0.01$ .**

(short vs. long) and the effect sizes across different adaptation phases. In the current study, we used odor stimuli, and the target stimuli (visual dots and sound beeps) were relatively emotionally neutral. Such an olfactory stimulus may be especially suitable for exploring the emotional response by itself because few attentional factors were involved (Gros et al., 2015). Hence, the confounding of attentional alertness induced by the stimuli themselves was minimized. We believe that attentional resources/engagements play an important role in time perception. The current study provides a good avenue to measure the attentional effect because the attentional and emotional factors (including the dimensions of valence and arousal) were separated from the inducers (odors) and the stimuli, thus making the investigations of the roles of attentional deployment and emotional states technically sound. Moreover, our results supported that the emotion induced from one sensory modality influenced time perception in another modality, indicating that there was crossmodal duration modulation (Shi et al., 2012).

Our results revealed a longer reproduction of time intervals than the actual time durations in the lavender condition as well as a shorter reproduction of time intervals in the jasmine and garlic conditions. These results could not be simply explained by arousal. According to the internal clock model (Gibbon et al., 1984), a high level of arousal would accelerate the rate of the pacemaker, increase the pulse of the accumulator, and lead to perceiving the duration as longer than it actually was. However, the present results contradict this prediction. This finding was because in previous studies (Tamm et al., 2014), perceived negative stimuli or threats (such as angry faces) led to the distribution of attentional resources between the tasks of time perception and emotion processing, which impaired time estimation for target events. The seemingly contradictory

results indicate that there might be other factors/mechanisms that modulate the bias in time perception. One possible reason is that the valence, rather than the arousal level of the stimuli, may play a major role in modulating time perception across the visual and auditory modalities when the inducers (odors) and target stimuli are separated. As we observed, lavender and jasmine were associated with a “positive” valence, while garlic was associated with a “negative” valence. The positive valence triggered more pulses, which led to a greater overestimation of the produced duration compared to the negative valence. This possibility is potentially weak because we found the opposite patterns with respect to the lavender and jasmine conditions (they had similar ratings for valence and arousal). Alternatively, the high arousal state in the garlic condition attracted attentional resources, which made the neutral stimuli comparatively less attended and decreased the reproduced duration (Tse et al., 2004). In the lavender condition, because the arousal level was low, more attention was directed to the neutral stimuli, and the perceived time intervals were longer. Even with the above arguments, one might consider that the special case of “jasmine” would not support the “attentional” accounts. We reserved the possibility that another dimension, such as personal preference (such as that for “jasmine”), would also attract the attentional resources for processing the time information of target stimuli.

Alternatively, one may also argue for a generally flexible framework for time estimation when different mechanisms, such as attention, arousal (valence), and other modulatory factors, come into play together. An attention mechanism may control the switch/gate, while an arousal mechanism may affect the rate of the pacemaker (Lake, 2016). Each mechanism may compete for general resources to play an important role in the process of time reproduction. For example, previous results have shown

that arousal is not the only main mechanism for time distortion because both arousal dependent time distortion (Droit-Volet et al., 2010) and arousal-independent time estimation (Schreuder et al., 2014) were reported. Moreover, attentional deployment might also act on the process of the pacemaker. Specifically, the distortion of time perception may increase according to whether attention is focused on time or on the signals. Furthermore, other factors, such as the gender of the participants (Grondin et al., 2015), anxiety (Bar-Haim et al., 2010), etc. could also modulate the effect size of emotional time distortions. Different mechanisms might counteract each other or neutralize the overall effects. Therefore, in the present study, the fact that we did not observe significant differences between each of the three odor conditions and the no-odor condition might be attributed to this consideration.

The attentional deployment in time perception was also supported by the evidence from the time course of the odor adaptation. In the present study, we found even longer reproductions of time intervals than the actual duration in the final block compared to the initial block of the experiment. An explanation of this finding is that with the passage of time, observers overcame through the influences of the odors, and the attentional resources were re-engaged to the target/neural stimuli (visual dots and auditory beeps). Therefore, we observed longer reproductions of the duration in the final block compared with the reproduced durations in the initial block. Thus, our results support that the function of the attentional mechanism varies over time because attention may be captured by emotional stimuli (Shi et al., 2012) or reduced after the repeated presentation of the emotional stimuli (Gros et al., 2015). It should be acknowledged that there are various adaption processes in the olfactory as well as in the limbic and cognitive systems during long-term smell exposure. Different odors and different dimensional properties of odors also have different time courses of adaptation. The short-spaced interval between the first and the final blocks in the present study could partly reduce the mixing effect of these factors. Nevertheless, further studies may investigate how the adaptation of odors affects time perception.

The attentional mechanism, however, is constrained by the actual length of the target duration. For the effects of odors as well as for their adaptation effect, we found unanimously that the effect sizes were larger in the “short” interval condition than in the “long” interval condition. Moreover, odor adaptation (i.e., the arousing states) influenced the perceived duration differently for the visual and auditory modalities for the short duration (1000 ms) but not for the long duration (4000 ms). As stated in the literature, the two ranges of time intervals (1000 and 4000 ms) may be different with respect to their underlying mechanisms (Poeppel, 1997). For example, the timing of a 1000 ms interval can be considered to be a relatively perceptual process, whereas the timing of a 4000 ms duration may involve higher cognitive functions and is usually referred to as time estimation (Fraisse, 1984; Poeppel, 1997). Moreover, for the short time duration of less than 1 s, the modality effect (i.e., the differences between the visual and auditory modalities) was easily observed. For the auditory signals, the rate of the internal clock was faster than that

for the visual signals, thereby inducing longer time perception in the auditory modality.

Although, the present study shed light on multiple mechanisms of emotional time perception, it is important to note that there are some limitations in our study. First, no low negative arousal odor was used in the present study, which made it impossible to interpret the ANOVA results for valence or arousal effects. Because most “negative” odors show high arousal patterns, we did not obtain satisfactory samples for the current study. However, the different and critical arousals and valences are present in the three odors used. To understand the different effect of valences or arousal effects of emotional stimuli, further studies should be conducted in the future. Second, as we noted earlier, for the subjective rating, no difference was found between the jasmine and lavender odors according to the self-report. In the future, to further examine the effect of arousal levels on time perception, the recordings of participants’ heart rate, skin conductance (Gros et al., 2015), blood flow and the other physiological indexes may be used to capture the objective evaluation of the “emotional” stimuli. Moreover, we hypothesized that there was a correspondence between the modulating effect of the attentional factor of the inducer and the target stimuli but the exact coupling of the two items requires further study. Finally, we did not find the significant difference between each of three odor conditions and the neutral condition, which might be due to the counteracting effect of the different mechanisms underlying emotional time perception. Further studies in which the emotional stimuli are presented only in an encoding phase and in which the reproduction phase is always neutral should be used (Noulhiane et al., 2007).

In sum, our results show that the perception of time duration is influenced by the presence of inducers (odors). Participants reproduced longer time intervals than the actual durations when exposed to the smell of lavender, and they reproduced shorter time intervals when exposed to the smells of jasmine and garlic. Our results indicated that a mixed mechanism, especially attentional deployment between the inducers (odors) and target stimuli, could largely account for the timing bias across different sensory modalities as well as the timing course of those biases. Those biases, however, were dependent on different target durations and showed that the processing of short and long intervals might use different mechanisms.

## AUTHOR CONTRIBUTIONS

ZY and TG designed the research; TG performed the research; TG, JW, and ZY analyzed the data; ZY and LC wrote the manuscript. All authors commented on and edited the manuscript. All authors approved it for publication.

## ACKNOWLEDGMENTS

The work was funded by grants from Natural Science Foundation of China (31470978) and the Center for Studies of the Hong Kong, Macao and Taiwan, Sun Yat-sen University.



## REFERENCES

- Allan, L. G., and Gibbon, J. (1991). Human bisection at the geometric mean. *Learn. Motiv.* 22, 39–58. doi: 10.1016/0023-9690(91)90016-2
- Bar-Haim, Y., Kerem, A., Lamy, D., and Zakay, D. (2010). When time slows down: the influence of threat on time perception in anxiety. *Cogn. Emot.* 24, 255–263. doi: 10.1080/02699930903387603
- Block, R. A., and Gruber, R. P. (2014). Time perception, attention, and memory: a selective review. *Acta Psychol.* 149, 129–133. doi: 10.1016/j.actpsy.2013.11.003
- Bobko, D. J., Thompson, J. G., and Schiffman, H. R. (1977). The perception of brief temporal intervals: power functions for auditory and visual stimulus intervals. *Perception* 6, 703–709. doi: 10.1068/p060703
- Brown, S. W. (1985). Time perception and attention: the effects of prospective versus retrospective paradigms and task demands on perceived duration. *Percept. Psychophys.* 38, 115–124. doi: 10.3758/BF03198848
- Bueti, D., Walsh, V., Frith, C., and Rees, G. (2008). Different brain circuits underlie motor and perceptual representations of temporal intervals. *J. Cogn. Neurosci.* 20, 204–214. doi: 10.1162/jocn.2008.20017
- Buhusi, C. V., and Meck, W. H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nat. Rev. Neurosci.* 6, 755–765. doi: 10.1038/nrn1764
- Diego, M. A., Jones, N. A., Field, T., Hernandez-Reif, M., Schanberg, S., Kuhn, C., et al. (1998). Aromatherapy positively affects mood, EEG patterns of alertness and math computations. *Int. J. Neurosci.* 96, 217–224. doi: 10.3109/00207459808986469
- Droit-Volet, S., Bigand, E., Ramos, D., and Bueno, J. L. O. (2010). Time flies with music whatever its emotional valence. *Acta Psychol. (Amst.)* 135, 226–232. doi: 10.1016/j.actpsy.2010.07.003
- Droit-Volet, S., and Gil, S. (2009). The time–emotion paradox. *Philos. Trans. R. Soc. B Biol. Sci.* 364, 1943–1953. doi: 10.1098/rstb.2009.0013
- Droit-Volet, S., and Meck, W. H. (2007). How emotions colour our perception of time. *Trends Cogn. Sci.* 11, 504–513. doi: 10.1016/j.tics.2007.09.008
- Fraisse, P. (1984). Perception and estimation of time. *Annu. Rev. Psychol.* 35, 1–37. doi: 10.1146/annurev.ps.35.020184.000245
- Gibbon, J., Church, R. M., and Meck, W. H. (1984). Scalar timing in memory. *Ann. N.Y. Acad. Sci.* 423, 52–77. doi: 10.1111/j.1749-6632.1984.tb23417.x
- Gil, S., and Droit-Volet, S. (2011). “Time flies in the presence of angry faces”... depending on the temporal task used! *Acta Psychol.* 136, 354–362. doi: 10.1016/j.actpsy.2010.12.010
- Goldstone, S., and Lhamon, W. T. (1974). Studies of auditory-visual differences in human time judgment: 1. Sounds are judged longer than lights. *Percept. Mot. Skills* 39, 63–82. doi: 10.2466/pms.1974.39.1.63
- Gros, A., Giroud, M., Bejot, Y., Rouaud, O., Guillemin, S., Eboulé, C. A., et al. (2015). A time estimation task as a possible measure of emotions: difference depending on the nature of the stimulus used. *Front. Behav. Neurosci.* 9:143. doi: 10.3389/fnbeh.2015.00143
- Gruber, R. P., and Block, R. A. (2013). The flow of time as a perceptual illusion. *J. Mind. Behav.* 34, 91–100.
- Grondin, S., Laflamme, V., Bienvenue, P., Labonte, K., and Roy, M. L. (2015). Sex effect in the temporal perception of faces expressing anger and shame. *Int. J. Comp. Psychol.* 28, 1–11.
- Kanai, R., Paffen, C. L., Hogendoorn, H., and Verstraten, F. A. (2006). Time dilation in dynamic visual display. *J. Vision* 6, 1421–1430. doi: 10.1167/16.12.8
- Kovar, K. A., Gropper, B., and Ammon, H. P. (1987). Blood levels of 1,8-cineole and locomotor activity of mice after inhalation and oral administration of rosemary oil. *Planta Med.* 53, 315–319. doi: 10.1055/s-2006-962725
- Lake, J. I. (2016). Recent advances in understanding emotion-driven temporal distortions. *Curr. Opin. Behav. Sci.* 8, 214–219. doi: 10.1016/j.cobeha.2016.02.009
- Lee, K. H., Seelam, K., and O'Brien, T. (2011). The relativity of time perception produced by facial emotion stimuli. *Cogn. Emot.* 25, 1471–1480. doi: 10.1080/02699931.2010.544455
- Moss, M., Cook, J., Wesnes, K., and Duckett, P. (2003). Aromas of rosemary and lavender essential oils differentially affect cognition and mood in healthy adults. *Int. J. Neurosci.* 113, 15–38. doi: 10.1080/00207450390161903
- Noulhiane, M., Mella, N., Samson, S., Ragot, R., and Pouthas, V. (2007). How emotional auditory stimuli modulate time perception. *Emotion* 7, 697–704. doi: 10.1037/1528-3542.7.4.697
- Ono, F., and Kitazawa, S. (2010). The effect of perceived motion-in-depth on time perception. *Cognition* 115, 140–146. doi: 10.1016/j.cognition.2009.12.006
- Penney, T. B., Gibbon, J., and Meck, W. H. (2000). Differential effects of auditory and visual signals on clock speed and temporal memory. *J. Exp. Psychol. Hum. Percept. Perform.* 26, 1770–1787. doi: 10.1037/0096-1523.26.6.1770
- Pöppel, E. (1997). A hierarchical model of temporal perception. *Trends Cogn. Sci.* 1, 56–61. doi: 10.1016/S1364-6613(97)01008-5
- Pöppel, E. (2004). Lost in time: a historical frame, elementary processing units and the 3-second window. *Acta Neurobiol. Exp.* 64, 295–301.
- Roberts, A., and Williams, J. M. G. (1992). The effect of olfactory stimulation on fluency, vividness of imagery and associated mood—a preliminary study. *Br. J. Med. Psychol.* 65, 197–199. doi: 10.1111/j.2044-8341.1992.tb01699.x
- Schreuder, E., Hoeksma, M. R., Smeets, M. A., and Semin, G. R. (2014). The effects of odor and body posture on perceived duration. *Front. Neurobot.* 8:6. doi: 10.3389/fnbot.2014.00006
- Schwartz, G. E., Whitehorn, D., Hernon, J. C., and Jones, M. (1986). Subjective and respiratory differences of fragrances: interactions with hedonics. *Psychophysiology* 23, 460.
- Shi, Z., Jia, L., and Müller, H. J. (2012). Modulation of tactile duration judgments by emotional pictures. *Front. Integr. Neurosci.* 6:24. doi: 10.3389/fnint.2012.00024
- Tamm, M., Uusberg, A. M., Allik, J., and Kreegipuu, K. (2014). Emotional modulation of attentional affects time perception: evidence from event-related potentials. *Acta Psychol.* 149, 148–156. doi: 10.1016/j.actpsy.2014.02.008
- Tipple, J. (2008). Negative emotionality influences the effects of emotion on time perception. *Emotion* 8, 127–131. doi: 10.1037/1528-3542.8.1.127
- Tremblay, S., and Fortin, C. (2003). Break expectancy in duration discrimination. *J. Exp. Psychol. Hum. Percept. Perform.* 29, 238–831. doi: 10.1037/0096-1523.29.4.823
- Tse, P. U., Intriligator, J., Rivest, J., and Cavanagh, P. (2004). Attention and the subjective expansion of time. *Percept. Psychophys.* 66, 1171–1189. doi: 10.3758/BF03196844
- Wearden, H., Edwards, M., Fakhri, A., and Percival, A. (1998). Why “sounds are judged longer than lights”: application of a model of the internal clock in humans. *Q. J. Exp. Psychol. B.* 51, 97–120. doi: 10.1080/713932672
- Welch, R. B., and Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychol. Bull.* 88, 638–667. doi: 10.1037/0033-2909.88.3.638
- Wittmann, M., and Paulus, M. P. (2008). Decision making, impulsivity and time perception. *Trends Cogn. Sci.* 12, 7–12. doi: 10.1016/j.tics.2007.10.004
- Zhang, X., and Zhou, X. (2007). Time perception of emotional events. *Prog. Nat. Sci.* 17, 150–153.

**Conflict of Interest Statement:** 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.

The handling Editor declared a current collaboration with one of the authors [LC] and states that the process nevertheless met the standards of a fair and objective review.

Copyright © 2016 Yue, Gao, Chen and Wu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# The duality of temporal encoding – the intrinsic and extrinsic representation of time

Ronen Golan<sup>1\*</sup> and Dan Zakay<sup>1,2</sup>

<sup>1</sup> School of Psychological Sciences, Tel Aviv University, Tel Aviv, Israel, <sup>2</sup> Psychology Department, Interdisciplinary Center Herzliya, Herzliya, Israel

## OPEN ACCESS

### Edited by:

Marc Wittmann,  
Institute for Frontier Areas  
of Psychology and Mental Health,  
Germany

### Reviewed by:

Evgeny Gutyrchik,  
Ludwig-Maximilians-Universität,  
Germany

Yuliya Zaytseva,  
Moscow Research Institute  
of Psychiatry, Russia

### \*Correspondence:

Ronen Golan,  
School of Psychological Sciences,  
Tel-Aviv University, Ramat-Aviv,  
Tel Aviv 69978, Israel  
golanronen@gmail.com

### Specialty section:

This article was submitted to  
Perception Science,  
a section of the journal  
Frontiers in Psychology

**Received:** 14 June 2015

**Accepted:** 12 August 2015

**Published:** 31 August 2015

### Citation:

Golan R and Zakay D (2015)  
The duality of temporal encoding –  
the intrinsic and extrinsic  
representation of time.  
Front. Psychol. 6:1288.  
doi: 10.3389/fpsyg.2015.01288

While time is well acknowledged for having a fundamental part in our perception, questions on how it is represented are still matters of great debate. One of the main issues in question is whether time is represented intrinsically at the neural level, or is it represented within dedicated brain regions. We used an fMRI block design to test if we can impose covert encoding of temporal features of faces and natural scenes stimuli within category selective neural populations by exposing subjects to four types of temporal variance, ranging from 0% up to 50% variance. We found a gradual increase in neural activation associated with the gradual increase in temporal variance within category selective areas. A second level analysis showed the same pattern of activations within known brain regions associated with time representation, such as the Cerebellum, the Caudate, and the Thalamus. We concluded that temporal features are integral to perception and are simultaneously represented within category selective regions and globally within dedicated regions. Our second conclusion, drawn from our covert procedure, is that time encoding, at its basic level, is an automated process that does not require attention allocated toward the temporal features nor does it require dedicated resources.

**Keywords:** time perception, temporal encoding, time representation, FFA, PPA, Cerebellum, Caudate, Thalamus

## Introduction

Encoding temporal information of our surrounding is a fundamental cognitive and neural process. A faithful representation of sensory information includes not just the WHAT and WHERE, but also the temporal characteristics of a stimulus or an event, i.e., the WHEN; however, we do not possess any temporal sensor. Moreover, opposed to other sensory information that can be turned on and off, temporal experience seems to be continuous and thus makes the task of tracing its nature much more challenging. Our goal in this study was to try and overcome this challenge and establish a methodology that can help us monitor, at the neural level, the automatic nature of temporal encoding.

While cognitive internal clocks models capture quite accurately our overt prospective time experience (Treisman, 1963; Church, 1984; Zakay, 1989; Zakay and Block, 1996), they still lack the mechanisms underlying our covert continuous temporal encoding and their relations to our conscious psychological time experience. Gaining a better understanding on the way temporal information is processed from its initial encoding stage up to its psychological experience stage

along with the neural substrates underlying it, may assist us in establishing a more comprehensive model of temporal processing.

Neural models of temporal encoding can be broadly divided into two categories – intrinsic vs. extrinsic or dedicated models. Intrinsic models relay on the idea that time is inherent to neural dynamics (i.e., oscillations, rhythmical or state-dependent models) and is represented locally in the brain. In this category we can find single cell models (Leon and Shadlen, 2003), where time is represented by the actual activity of a specific neuron: either excitatory (i.e., more is more; higher activity means longer duration; Pariyadath and Eagleman, 2007; Eagleman, 2008), or inhibitory (Constantinidis et al., 2002) by inhibiting a response for a specified duration, as well as “State Dependent Models” where time is represented based on the general properties of a specific neural network (Buonomano and Merzenich, 1995; Buonomano, 2000; Karmarkar and Buonomano, 2007). Extrinsic or dedicated models focus on central specialized time keeping mechanism, such as central internal clock models (Meck, 1996, 2005; Matell and Meck, 2000; Hofstötter et al., 2002; Mauk and Buonomano, 2004) where time is globally represented based on an oscillating unit; or, network models involving several centers in the brain such as the Cerebellum-SMA-Basal Ganglia circuit (Ferrandez et al., 2003) along with scattered representation of time, where several internal clocks represent time for different modalities (Goldstone and Lhamon, 1974; Lhamon and Goldstone, 1974; N'Diaye et al., 2004).

Naturally there are pros and cons for each of these approaches. As expressed by Ivry and Schlerf (2008), dedicated models have difficulties in accounting for impaired encoding of time as a result of modulations in neural activity. At the same time intrinsic models suffer from poor explanatory power when accounting for cross modality effects on time perception (Warm et al., 1975; Roberts, 1982). Moreover, intrinsic models will have difficulties to account for global effects on time perception, like attention, while extrinsic models will have difficulties explaining local representation of time of specific types of stimuli in category selective areas.

Either intrinsic or extrinsic, current models seem to lack a more radical approach, namely that temporal characteristics are integral to stimulus processing and should be represented within category or feature selective neural populations.

A related issue in the study of time perception and temporal representation is whether time encoding and perception are automated, continuous, pre-attentive processes or do they require the allocation of attention or dedicated resources to a well-defined duration or interval; in other words, the overt vs. the covert encoding of time. Some studies suggested that implicit timing has distinct mechanisms (Coull and Nobre, 2008) while other studies (Praagstra et al., 2006) suggested that implicit and explicit timing relies on common mechanisms. Nevertheless, the mere fact that time can be represented implicitly (regardless of the mechanisms underlying it) suggest that time encoding has an automatic pre-attentive characteristic that should be addressed.

In the current study, our main challenge was to establish a method for tracking covert temporal processing at the neural

level. Based on the assumption that time is an integral part of perception (Zakay et al., 2014), we expected temporal encoding to be represented within sensory modules, associated with the representation of the stimulus, rather than solely associated with either mere neural dynamics or distinct generic regions. More specifically, we expected category selective areas that typically encode information about the shape of visual stimuli to also encode its temporal information. Consequently, our two main goals in this study were to test whether temporal representation is an integral part of the stimulus representation occurring within category selective brain regions that elicit a selective response to specific stimuli; and to test if the temporal encoding is an automatic process occurring without allocating attention toward the temporal characteristics of the stimulus and without involving any motor reaction or motor planning.

In order to achieve these goals, functional MRI was used to investigate both the continuous property of time, namely the encoding of time without allocating dedicated resources for its processing; and the dual simultaneous representation of time, both intrinsically and extrinsically. We based our method on the novelty-habituation effect of neural population behavior. Perceiving neurons are excited when presented with a novel stimuli; a repetitive presentation of a stimulus will instigate an habituation process which is associated with neuronal inhibitory processes. The presentation of a novel stimulus will reinstate neuronal excitation (i.e., dishabituation). The processes described above suggest some sort of a feature comparison mechanism as per novelty detection. Indeed, Sokolov (1990) suggested a feature neural comparison model which was supported by several studies (see Siddle, 1991; Ben-Shakhar et al., 2000; Ben-Shakhar and Gati, 2003). Feature mismatch during this comparison procedure is assumed to be the basis for neuronal excitation.

A step toward adopting this novelty/habituation approach in fMRI was practiced by Grill-Spector et al. (1999) using a method called fMR-Adaptation (fMR-A; for an overview see Grill-Spector and Malach, 2001 and Grill-Spector et al., 2006). Grill-Spector and Malach (2001) repeatedly exposed subjects to visual objects stimuli in order to study invariant object properties (e.g., rotation, illumination) in high-order object areas [Lateral Occipital Complex (LOC)]. These studies demonstrated that by examining which object properties showed release from adaptation (i.e., dishabituation) the nature of representation of different object features in LOC (e.g., invariance to object rotation) would surface.

We believe that fMR-A can be applied to the study of duration encoding. Specifically, we aimed to apply the fMR-A technique in order to find brain regions that show dishabituation to temporal information of visual stimuli. We estimated that such a procedure could be used to localize brain areas that are sensitive to temporal information and therefore are involved in the encoding stages of time perception. Consequently, we designed a covert procedure where subjects were not informed of the temporal settings of the experiment, while engaged in a non-motor, non-temporal task. Moreover, we selected distinct brain regions that typically encode information about the shape of visual stimuli and asked if temporal encoding is performed within these regions. More

specifically, we functionally localized the face-selective areas in the inferior occipital lobe (occipital face area – OFA) and in the fusiform gyrus (fusiform face area – FFA) as well as scene-selective areas such as the parahippocampal place area (PPA) and the transverse occipital sulcus (TOS). The reason for selecting these regions is that they are well-defined and easy to functionally localize.

In line with the neural behavior where repetitive exposure to the same feature or stimulus decreases the neural activation of the subpopulation representing it (Grill-Spector et al., 2006; Krekelberg et al., 2006) we expect that neural activation of subpopulations specializing in representing a specific stimulus will be positively correlated to the variance in stimulus features, including its temporal features. That is to say, the higher the variance between repetitive stimuli – the higher the activation of the neural subpopulation would be. Thus, we hypothesized that if the specific regions mentioned above (FFA, OFA, PPA, and TOS) encode the temporal features of their category stimuli, they should be sensitive to variance in the exposure duration of these stimuli. More specifically, we expected that the higher the variance in exposure duration – the larger the elicited activation of the specific region should be.

## Materials and Methods

### Participants

Fifteen healthy volunteers (12 females and 3 males; mean age = 28.47 years, SD = 4.37) participated in the Experiment. All subjects were either undergraduate students or post graduates

with advanced degrees (i.e., MAs, PhDs, MD). All subjects gave informed consent for participation in the study, which was approved by the ethics committee of the Tel-Aviv Sourasky Medical Center.

### Stimuli

All visual stimuli were grayscale. Functional localizer images were miscellaneous 300 pixels × 300 pixels photographs of 80 different faces, 80 different types of objects (e.g., ball, apple, barrel) and 80 different types of natural scenes (e.g., houses, landscapes).

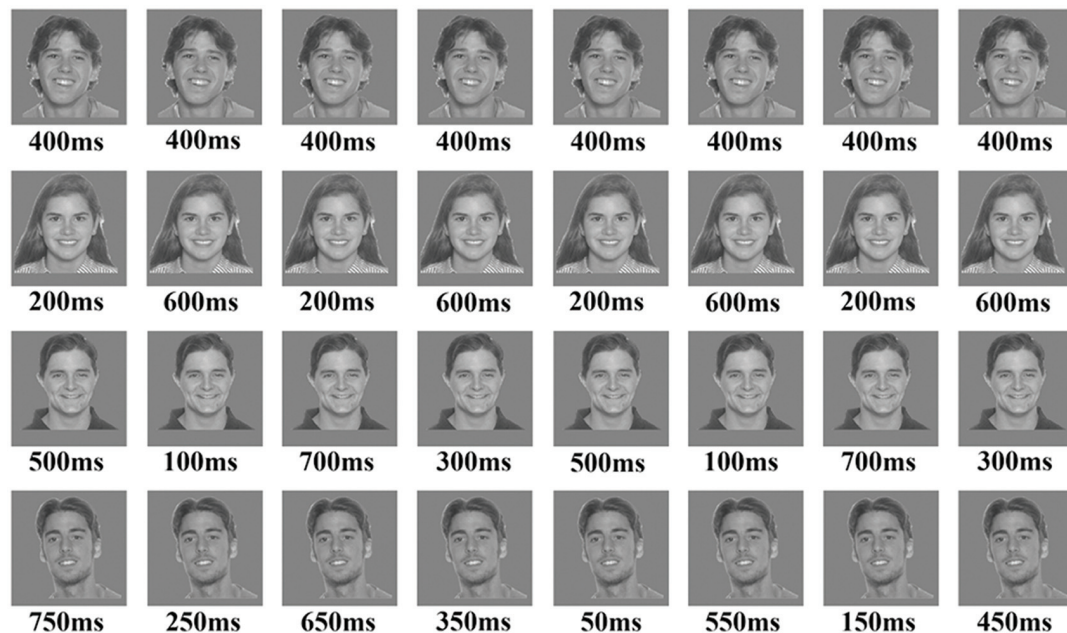
Stimuli used in the temporal condition scans were four different photographs of faces (**Figure 1**) and four different photographs of houses with an average intensity of 160. Background was grayscale with an intensity of 160 to match the average intensity of the stimuli. Faces and houses images were randomly selected from the images used in the localizer scans; all faces images had the same expression. An altering colored fixation point of 8 pixels × 8 pixels was presented in the center of the images, using Matlab 7 (Psychtoolbox, Brainard, 1997).

Stimuli were presented with Matlab 7 (Psychtoolbox, Brainard, 1997) and were projected onto a screen located at the back of the scanner through a projector. Subjects viewed the stimuli through a mirror that was placed on the upper part of an eight channel head coil in front of their eyes.

### Procedure

#### Functional Localizer Scans

A block-design functional localizer was used to identify regions of interest (ROI) of two categories, Faces and Scenes, using three



**FIGURE 1 | The figure presents an example of the four blocks in the temporal condition using face stimuli.** The upper row is an example of the baseline block with 0% variance, while the fourth row is an example of block 4

with 50% variance in durations. The parameters for each block consisted of the following: (a) Number of stimuli in a block was 16; (b) ISI was 350 ms; (c) block duration was 12-s; (d) total exposure to stimulus in all blocks was 6400 ms.



stimulus classes: Faces, Scenes, and Objects. Functional localizer scans consisted of four blocks (16-s each) of each stimuli category (each condition repeated four times in each scan) and five blocks (16-s each) of a baseline fixation point. For each block, 20 images form a single stimulus class were presented (300-ms per image, with 500-ms interstimulus interval); with an addition of a 12-s dummy block, each localizer scan lasted for 4-min and 44-s. In order to ensure general vigilance, subjects were instructed to memorize all consecutive identical images (1-back task). At the end of each scan subjects were asked to report their findings.

### Temporal Conditions Scans

For the temporal scans we used a block-design, where each scan consisted of four condition blocks (12-s each) presented twice (total of eight blocks), and nine baseline fixation blocks (12-s each). In each block subjects were exposed to 16 repetitions of the exact same stimulus of the same category (either a face or a house). For each condition block, stimulus duration had a distinct degree of variance (i.e., 0, 12.5, 25, and 50%), while the ISI between stimuli and the total exposure to a stimulus within a block remained constant across all blocks. In order to make sure that the effect is not duration dependent, we used two sets of durations, as well as two types of ordering (i.e., the first duration in a block could be either the relatively long duration or the relatively short one). Thus, eight subjects were exposed to the first set, and seven subjects to the second set.

In the 0% variance condition, each repetition had the same duration (either 400-ms for some subjects; or 500-ms for others). In the 12.5% variance condition we used two types of durations (either 600 and 200-ms with eight repetitions each; or 200 and 800-ms; see **Figure 1**). For the 25% variance, we used four types of durations with four repetitions each (either 100, 300, 500, and 700-ms; or 150, 300, 700, and 850-ms). And finally, for the 50% variance condition we used eight types of durations with two repetitions each (either 50, 150, 250, and 350-ms, 450, 550, 650, and 750-ms; or 100, 250, 400, and 450-ms, 550, 600, 750, and 900-ms). Within the first set of durations ISI was 350-ms and total exposure time within a block was 6400-ms for all conditions. Within the second set of durations ISI was 250-ms and the total exposure time within all conditions was 8000-ms.

In total there were four scans using faces images and four scans using houses images. Conditions were counterbalanced both for order and images so that a specific image will appear across all types of conditions.

A colored fixation dot was presented over each image. Participants were instructed to detect a distinct pattern of the colored fixation point (e.g., a sequence of two consecutive blue dots). To avoid any motor reaction during the experiment, at the end of each scan subjects were instructed to report in which of the images the distinct sequence appeared more frequently.

## MRI

### Data Acquisition

MRI data was collected in a 3T GE MRI scanner. Echo planar imaging sequence was used to collect fMRI data with the

following parameters: TR = 2-s, TE = 35-ms, flip angle: 90°, 34 slices per TR, slice thickness: 4 mm no gap, matrix 64 × 64, and FOV 256-mm.

### Data Analysis

fMRI data analysis was conducted using statistical parametric mapping (SPM2<sup>1</sup>). Images acquired during the first 12-s of each scan were discarded. Preprocessing of EPI images included slice timing correction, realignment, normalization to a standard template [Montreal Neurological Institute (MNI), voxel size 3 × 3 × 3], and spatial smoothing with an 5 mm × 5 mm × 5 mm full-width at half-maximum (FWHM) Gaussian kernel.

### Functional Localizer Data Analysis

We used data from the localizer scans to identify regions dedicated for processing of faces and scenes for each subject independently. The FFA and OFA for face selective area were defined by using a Faces > Objects contrast. For the scenes selective area a Scene > Object contrast was used to define the PPA and the TOS.

### Temporal Conditions Scans within Category Selective Areas Data Analysis

A general linear model was estimated for each individual subject using SPM2, Finite Impulse Responses (FIR) and Fitted event time courses were extracted from the predefined ROIs using the MarsBar Toolbox for SPM2<sup>2</sup> and imported into MATLAB R2010a and SPSS21 for statistical analysis. For the statistical analysis we used the maximum value of the FIR signal percent change from each time course and the two highest values surrounding it (a total of 6-s out of a 12-s block).

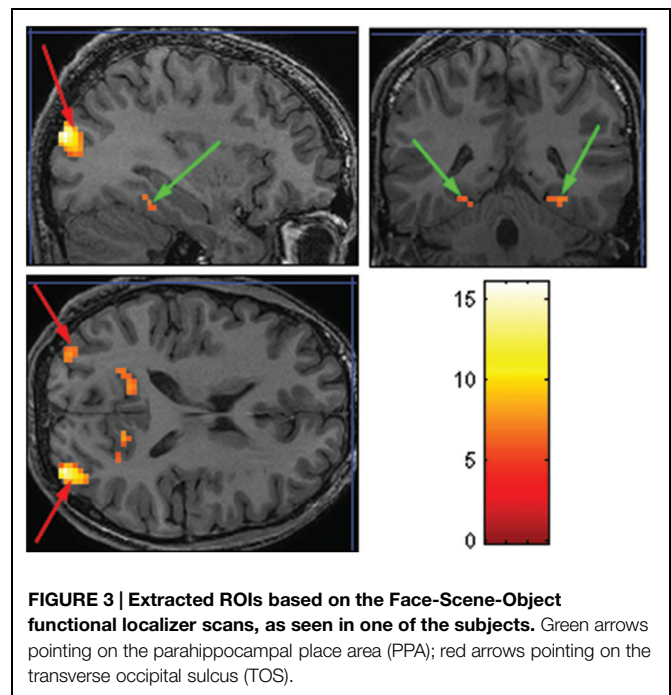
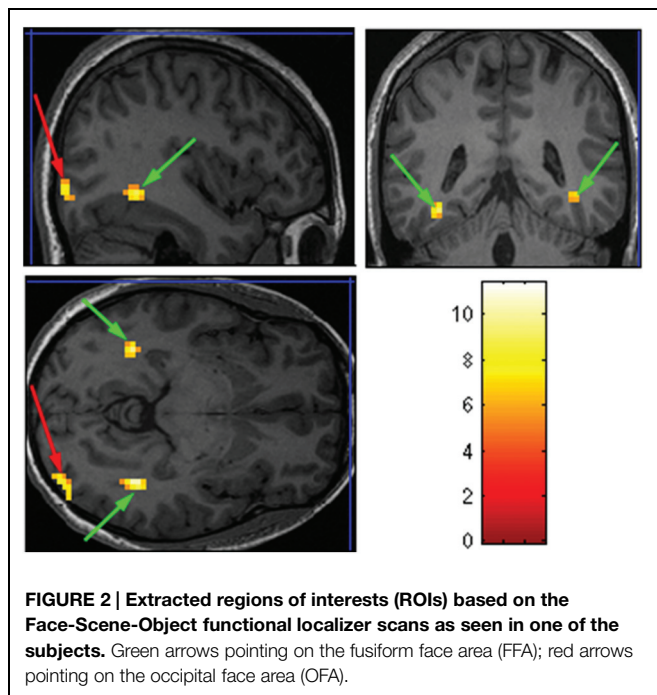
### Whole Brain Second Level Analysis for Temporal Conditions Scans

A general linear model was estimated for each individual subject. A whole brain analysis was performed for each subject using a parametric contrast (i.e., −3, −1, 1, 3 for the temporal conditions, respectively), while disregarding the distinction between faces stimuli and houses stimuli. As a result, for the second level analysis each condition included data from 16 blocks rather than 8 for the category selective areas.

Contrasts from each subject were used for a second level analysis (i.e., “basic models” *t*-Test in SPM2). Based on the second level whole brain analysis we extracted common ROIs of significant activations ( $p = 0.0001$  uncorrected). FIR and Fitted event Time courses were extracted from each individual subject, based on extracted ROIs using the MarsBar Toolbox for SPM2 and imported into MATLAB R2010a and SPSS21 for statistical analysis. For the statistical analysis we used the maximum value of the FIR signal percent change from each time course and the two highest values surrounding it (a total of 6-s out of a 12-s block).

<sup>1</sup><http://www.fil.ion.ucl.ac.uk/spm/software/spm2/>

<sup>2</sup><http://marsbar.sourceforge.net/>



## Results

### Face/Scene Localizer

The purpose of the Face-Scene localizer was to extract specific ROIs and to test how these ROIs respond to the temporal manipulation. We used the Faces > Objects contrast ( $p = 0.001$  uncorrected) to identify the FFA (Figure 2 green arrows) and OFA (Figure 2 red arrows) and the Scene > Object contrast ( $p = 0.001$  uncorrected) to identify the PPA (Figure 3 green arrow) and TOS (Figure 3 red arrow). Based on our findings and on previous studies (Kanwisher et al., 1997; Loffler et al., 2005; Rotshtein et al., 2005), we focused our analysis on activations in the right hemisphere (i.e., rFFA, rOFA, rPPA, and rTOS), which elicited stronger activations than in the left hemisphere (for example only 6 out of the 15 Ss showed left OFA activation).

Out of our 15 subjects one subject did not show any specific activations to neither faces nor houses stimuli. Moreover, one subject did not show activations only in rOFA, another subject had no activations only in rPPA and a fourth subject had no activations in rTOS. Thus for the analysis of temporal conditions within category selective ROIs we used data from 14 Ss for rFFA and from 13 Ss for rOFA, rPPA, and rTOS. However, for the second level analysis where no pre-defined ROIs were required, data from all participants was included.

### Temporal Conditions within Category Selective Areas

After identifying face and scene selective areas for each subject (e.g., rFFA, rOFA, rPPA, and rTOS), we extracted the time courses for each temporal condition from each ROI for each subject.

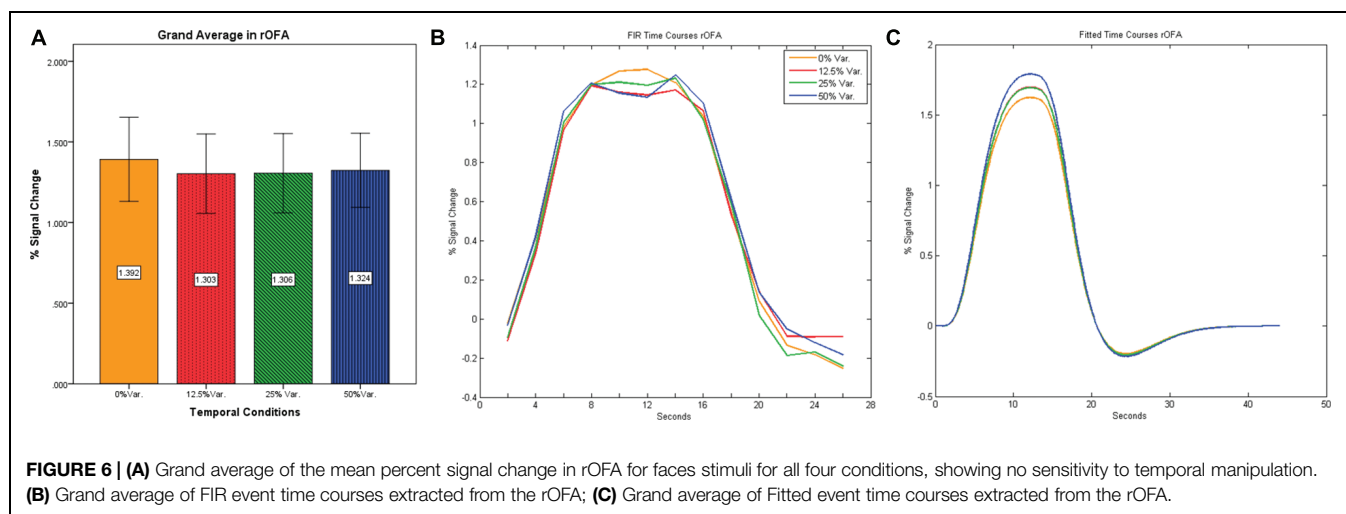
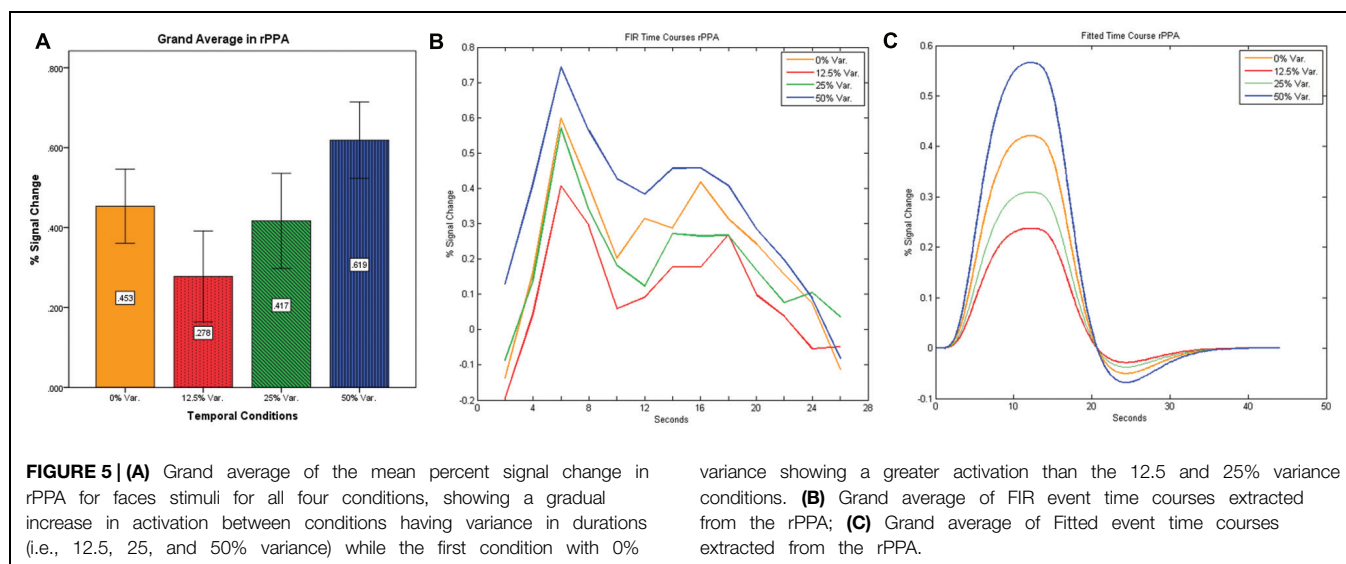
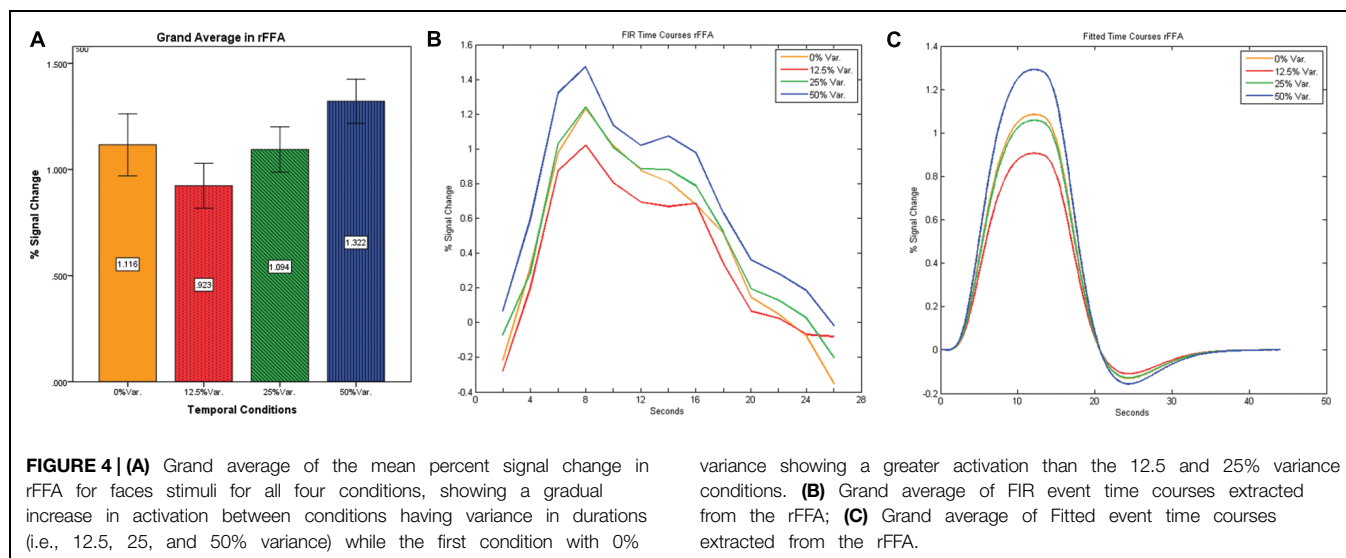
As can be seen in Figures 4 and 5 we found that the 0% variance condition yielded a lower activation than the 50%

variance condition both in rFFA [ $t(82) = -2.29, p = 0.025$ ] and rPPA [ $t(76) = -2.48, p = 0.015$ ]. Moreover, a gradual increase in mean activations appeared between the second, third, and fourth conditions in both the rFFA and the rPPA. The first condition yielded a higher activation than the second condition in both ROIs. A one-way ANOVA test over all four conditions revealed a significant main effect of the difference in mean activations for the rFFA with faces stimuli [ $F(3,164) = 7.77, p < 0.00007$ ] and for the rPPA with houses stimuli [ $F(3,152) = 7.01, p = 0.0002$ ]. A linear trend test using contrast coefficients of  $[-3, -1, 1, 3]$  yielded a significant effect in both rFFA [ $t(164) = 3.01, p = 0.003$ ] and rPPA [ $t(152) = 2.68, p = 0.008$ ].

In contrast to the rFFA and rPPA, no effect was found in rOFA [ $F(3,152) = 0.114, p = 0.952$ ] nor in rTOS [ $F(3,152) = 0.595, p = 0.62$ ] (see Figures 6 and 7). Consequently no significant linear trend was found within these ROIs, i.e., rOFA – [ $t(152) = -0.37, p = 0.714$ ]; rTOS – [ $t(152) = 0.45, p = 0.653$ ].

### Second Level Analysis of the Temporal Conditions – Whole Brain Analysis

We conducted a second level analysis including all 15 subjects while disregarding stimulus type (i.e., faces or houses) and addressing only the four temporal conditions differing in variance of the duration exposure of the stimulus to extract common ROIs sensitive to the temporal variance manipulation. Analysis yielded four bi-lateral distinct ROIs. MNI coordinates were identified based on Talairach human brain mapping, returning the following four regions (Figure 8): (1) Right and Left Parahippocampal Gyrus and; (2) the right and left Thalamus; (3) right and left Caudate; and (4) right and left Cerebellum with specificity to the Pyramis, Inferior Semilunar lobule, and Culmen. (see Table 1 for details of ROIs and MNI Coordinates).



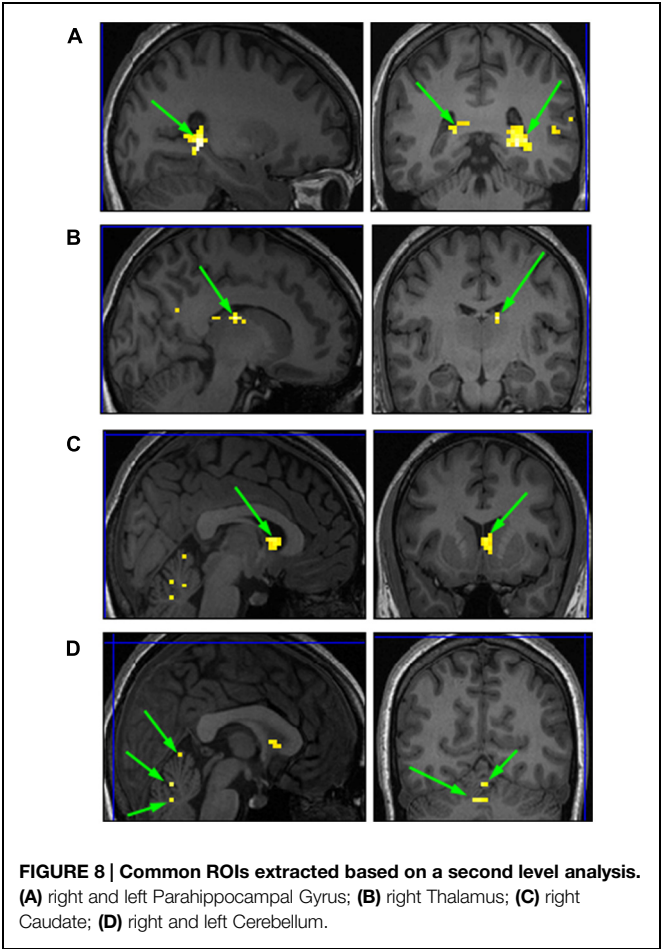
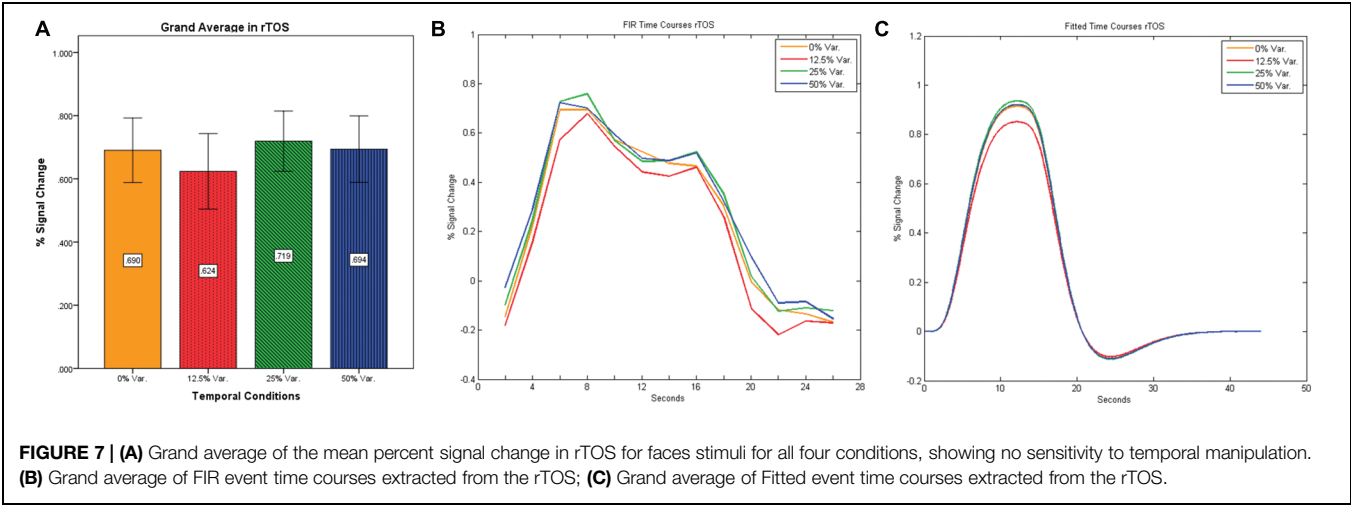


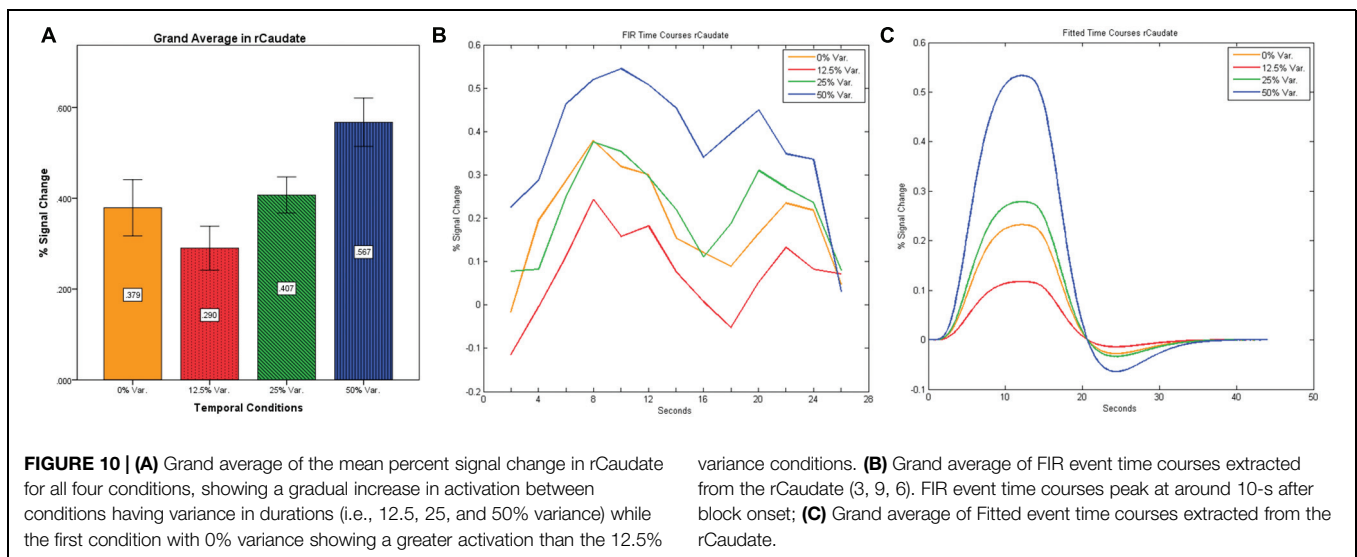
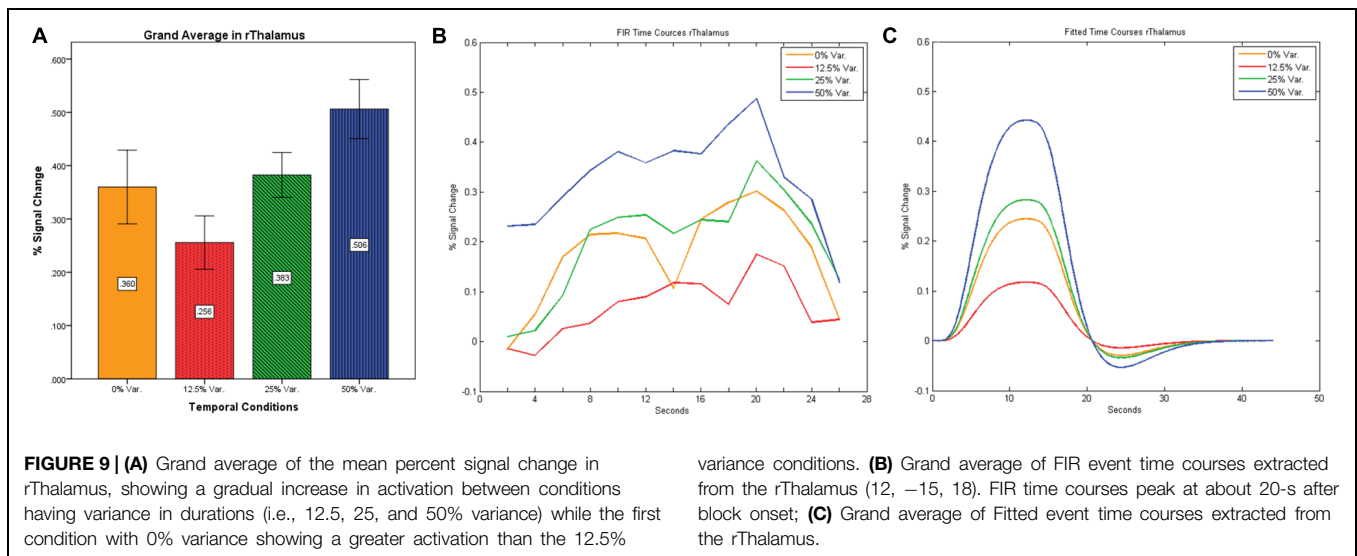
TABLE 1   Common regions of interests (ROIs) based on a second level analysis.	
Montreal Neurological Institute coordinates (x, y, z)	Talairach mapping
(27, -42, -6)	Right Parahippocampal Gyrus
(-18, -45, 3)	Left Parahippocampal Gyrus
(3, 9, 6)	Right Caudate Body/Head
(-3, 15, 6)	Left Caudate Body/Head
(24, -33, 18)	Right Caudate Tail
(-6, 0, 12)	Left Caudate Body
(12, -15, 18)	Right Thalamus
(-9, -24, 12)	Left Thalamus – Pulvinar
(0, -63, -24)	Left/Right Cerebellum – Pyramis
(-3, -63, -36)	Left Cerebellum – Inferior Semi-Lunar Lobule
(-3, -57, -3)	Left Cerebellum – Culmen

subjects yielded the exact same pattern as in the FFA and PPA, that is, a gradual increase in activation between the 12.5, 25, and 50% variance conditions, while the 0% condition yielded a higher activation the 12.5% and the 25% variance conditions (see **Figures 9–12** showing mean percent signal change, FIR time courses and Fitted time courses for the right hemisphere ROIs). A one-way ANOVA test and a linear trend test using contrast coefficients of (−3, −1, 1, 3) over all four conditions, revealed a significant main effect as well as a significant linear trend, for all ROIs (see **Table 2**).

### Discussion and Conclusion

Our main finding in this study was the apparent representation of time in object category selective areas. As results indicate, category selective areas that are typically associated with the representation of shapes (i.e., faces or houses) are also sensitive to the variance in the exposure duration of these shapes. Consequently we conclude that temporal encoding is an integral part of perception.





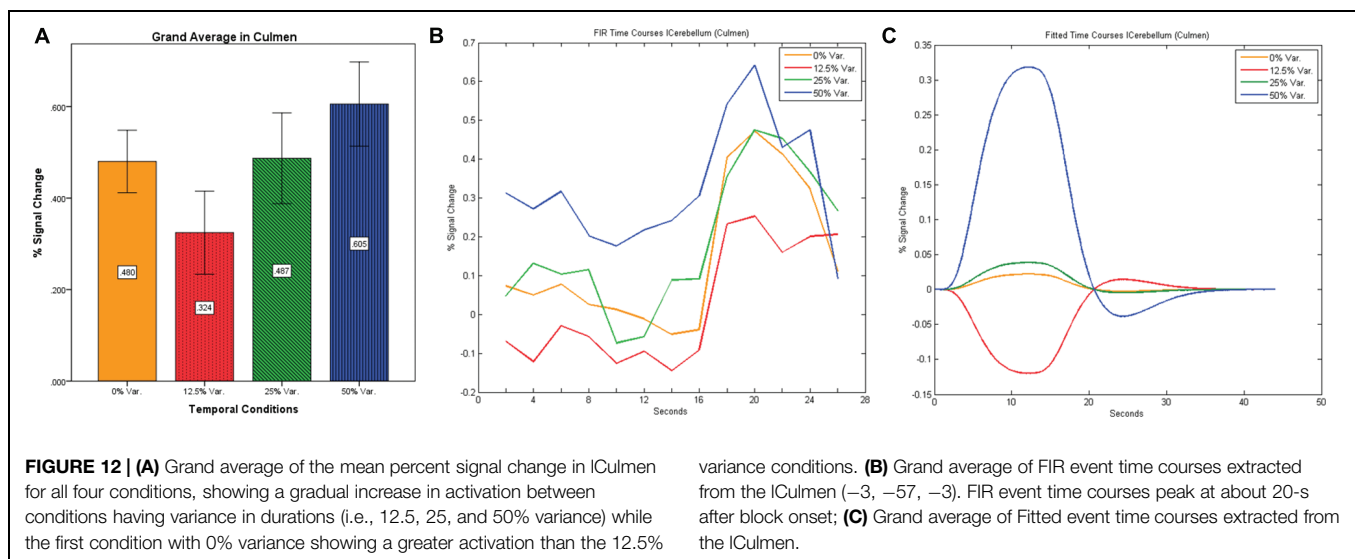
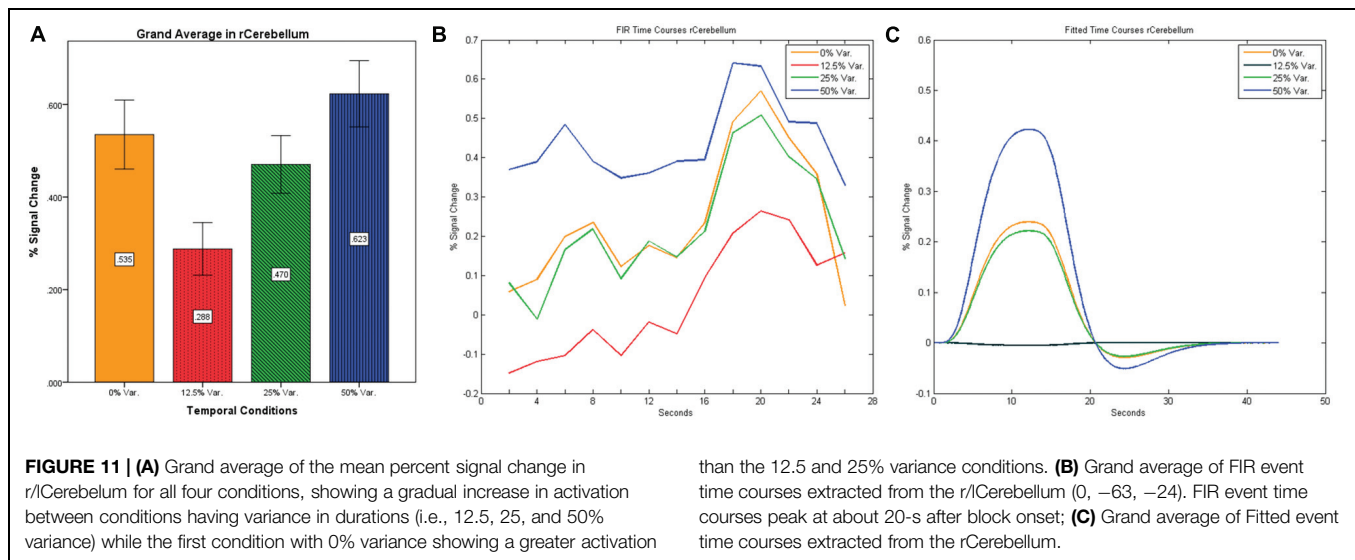
Moreover, it appears that this sensitivity to variations in the durations of the stimuli does not appear in occipital (dorsal) regions such as the OFA and TOS, and thus should be assigned specifically to ventral regions (i.e., FFA and PPA). This conclusion also suggests that the effect found is not a general attention effect, or a global effect to variance, but rather reflects the specific sensitivity of these regions to variance in durations.

Our second finding relates to global effects associated with our temporal manipulation. Findings suggest that when disregarding stimuli specificity (faces vs. houses) and testing globally only for effects based on the amount of variance in duration, we find four distinct areas that seem to be sensitive to time (or at least to the duration variance of stimuli): (1) The Parahippocampal Gyrus; (2) the Thalamus; (3) the Basal Ganglia with specificity to the Caudate; and (4) the Cerebellum having three inner distinct sub-regions, the Culmen among them. While the involvement of the

Parahippocampal Gyrus is directly related to the nature to the experiment, being sensitive to the specific stimuli presented, the involvements of the Thalamus, Caudate, and Cerebellum were less predictable.

A conclusion that can be drawn from the fact that our subjects were not informed of the temporal nature of the experiment and were engaged in a non-temporal task is that the sensitivity to duration variance is based on automatic processes which do not require attention or dedicated cognitive resources to process durations.

As results show, either locally within category selective regions or globally in the Cerebellum, Caudate, and Thalamus, the neural activation pattern was the same. While we expected an inclining gradient where the 0% variance condition will yield the minimum amount of activation while the 50% variance condition will yield the maximum neural activation, we found that this inclining gradient appears only within conditions where variance exist (i.e., 12.5, 25, and 50% variance) while 0% conditions yielded



**TABLE 2 | One-way ANOVA and linear trend analysis.**

MNI coordinates (x, y, z)	Talairach mapping	One-way ANOVA	Linear trend test Contrast coefficients: (–3, –1, 1, 3)
(3, 9, 6)	Right Caudate Body/Head	$F(3,176) = 20.14, p < 0.0001$	$t(176) = 5.92, p < 0.0001$
(–3, 15, 6)	Left Caudate Body/Head	$F(3,176) = 19.55, p < 0.0001$	$t(176) = 5.48, p < 0.0001$
(24, –33, 18)	Right Caudate Tail	$F(3,176) = 28.14, p < 0.0001$	$t(176) = 5.70, p < 0.0001$
(–6, 0, 12)	Left Caudate Body	$F(3,176) = 13.11, p < 0.0001$	$t(176) = 4.56, p < 0.0001$
(12, –15, 18)	Right Thalamus	$F(3,176) = 13.96, p < 0.0001$	$t(176) = 4.59, p < 0.0001$
(–9, –24, 12)	Left Thalamus – Pulvinar	$F(3,176) = 17.42, p < 0.0001$	$t(176) = 5.61, p < 0.0001$
(0, –63, –24)	Left/Right Cerebellum – Pyramis	$F(3,176) = 18.11, p < 0.0001$	$t(176) = 2.99, p = 0.003$
(–3, –63, –36)	Left Cerebellum – Inferior Semi-Lunar Lobule	$F(3,176) = 20.12, p < 0.0001$	$t(176) = 3.15, p = 0.002$
(–3, –57, –3)	Left Cerebellum – Culmen	$F(3,176) = 6.81, p = 0.0002$	$t(176) = 2.73, p = 0.007$

on average, a higher activation than the 12.5% and in several cases than the 25% variance conditions. We suspect that the higher activation of the 0% variance condition may lay on the fact that ecologically it differed from the other conditions as

being the sole condition with no variance and may resemble the comparison of responses to shades of the red color (i.e., variance conditions) vs. the responses to the color blue (i.e., 0% condition).

## Implications on Duration Encoding and Time Perception Model

The present experiment contributes to the study of intrinsic and dedicated models of temporal representation. Results in this experiment suggest that on the one hand when looking for local representation of time, one should look within category selective brain regions that are typically associated with the encoding or representation of the specific type of stimulus at hand; while on the other hand, simultaneously, time is also represented globally. Based on the paradigm used in this study, this dual representation is not related to a temporal task or attentional resources allocated to temporal features.

As far as we know, this study is the first to report on temporal representation within the FFA and PPA with stimuli specificity. However, with respect to the Thalamus, Cerebellum, and the Caudate, this study seems to be in line with previous findings with two main exceptions: the first is that in the preset study no temporal or motor task were involved; and the second is that we used natural visual images of faces and houses stimuli, and thus the involvement of the Cerebellum, the Basal-Ganglia, and Thalamus in processing, or at least in being sensitive to variance in one of the visual stimuli properties (i.e., duration), was unexpected. The Cerebellum and the Basal Ganglia are common regions found to be associated with time processing. The Thalamus, however, is less common. Meck (1996) and Matell and Meck (2004) suggested that the Thalamus might be part of a time keeping circuit involving the Basal-Ganglia, which by itself is part of a larger time circuit involving the Cerebellum. Moreover, based on evidence showing high variability in time estimation in patients with lesions in Basal Ganglia and Thalamus, Gibbon et al. (1997) suggested that the Thalamus might have some sort of a regulatory function over information coming from the Cerebellum, or play a part in mediating information which is processed in the Basal Ganglia and transferred to the Putamen. Lee et al. (2007) suggested a more specific circuit of subseconds

of time perception consisting of a Cerebral-Thalamus-Basal Ganglia-Cerebellum circuit. Several other studies report and suggest the involvement of the Thalamus in a loop network or circuit pertaining to time perception or time encoding (Ferrandez et al., 2003; Teki et al., 2011). The shared facet between most of these studies is that the Thalamus mediates information from the cortex to the inner ganglia (i.e., Basal Ganglia), or to deeper structures (i.e., the Cerebellum). In some models, the SMA and preSMA are also involved (Rao et al., 1997; Macar et al., 2004; Buhusi and Meck, 2005).

Our findings suggest that there should be a distinction in the way time is represented between the Cerebellum and the Thalamus on the one hand, and the Basal-Ganglia on the other. As can be seen in **Figure 10**, time courses in the Basal Ganglia peaks at about 10-s after block onset. This finding is in line with our findings of the way time courses in category selective areas behave under this specific manipulation. However, time courses in the Cerebellum and the Thalamus differ significantly from that pattern. As can be seen (see **Figures 9, 11, and 12**), time courses in these regions peak at about 20-s after block onset (during fixation time) suggesting a secondary role in temporal encoding that may be based on temporal information processed in the Basal-Ganglia. These findings are supported by Rao et al. (2001), which presents an Event-Related fMRI experiment also showing that temporal processing in the Basal-Ganglia occurs relatively early with respect to the Cerebellum which they assigned to the Cerebellum involvement in the process of timing rather than to the encoding of explicit timing.

Naturally, additional studies need to further investigate this issue, and better understand the nature of the delayed peak as well as the nature of the relationship between the Cerebellum, the Thalamus, and the Basal-Ganglia. Moreover, in order to generalize our findings relating to the representation or encoding of durations in category or feature selective regions, further exploration is needed.

## References

- Ben-Shakhar, G., and Gati, I. (2003). The effects of serial position and frequency of presentation of common stimulus features on orienting response reinstatement. *Psychophysiology* 40, 139–145. doi: 10.1111/1469-8986.00014
- Ben-Shakhar, G., Gati, I., Ben-Bassat, N., and Spiner, G. (2000). Orienting response reinstatement and dishabituation: effects of substituting, adding, and deleting components of nonsignificant stimuli. *Psychophysiology* 37, 102–110. doi: 10.1111/1469-8986.3710102
- Brainard, D. H. (1997). The psychophysics toolbox. *Spat. Vis.* 10, 433–436. doi: 10.1163/156856897X00357
- Buhusi, C. V., and Meck, W. H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nat. Rev. Neurosci.* 6, 755–765. doi: 10.1038/nrn1764
- Buonomano, D. V. (2000). Decoding temporal information: a model based on short-term synaptic plasticity. *J. Neurosci.* 20, 1129–1141.
- Buonomano, D. V., and Merzenich, M. M. (1995). Temporal information transformed into a spatial code by a neural network with realistic properties. *Science* 267, 1028–1030. doi: 10.1126/science.7863330
- Church, R. M. (1984). Properties of the internal clock. *Ann. N. Y. Acad. Sci.* 423, 566–582. doi: 10.1111/j.1749-6632.1984.tb23459.x
- Constantinidis, C., Williams, G. V., and Goldman-Rakic, P. S. (2002). A role for inhibition in shaping the temporal flow of information in prefrontal cortex. *Nat. Neurosci.* 5, 175–180. doi: 10.1038/nn799
- Coull, J. T., and Nobre, A. C. (2008). Dissociating explicit timing from temporal expectation with fMRI. *Curr. Opin. Neurobiol.* 18, 137–144. doi: 10.1016/j.conb.2008.07.011
- Eagelman, D. (2008). Human time perception and its illusions. *Curr. Opin. Neurobiol.* 18, 131–136. doi: 10.1016/j.conb.2008.06.002
- Ferrandez, A. M., Hugueville, L., Lehericy, S., Poline, J. B., Marsault, C., and Pouthas, V. (2003). Basal ganglia and supplementary motor area subsecond duration perception: an fMRI study. *Neuroimage* 19, 1532–1544. doi: 10.1016/S1053-8119(03)00159-159
- Gibbon, J., Malapani, C., Dale, C. L., and Gallistel, C. R. (1997). Toward a neurobiology of temporal cognition: advances and challenges. *Curr. Opin. Neurobiol.* 7, 170–184. doi: 10.1016/S0959-4388(97)80005-80000
- Goldstone, S., and Lhamon, W. T. (1974). Studies of auditory-visual differences in human time judgment: 1. Sounds are judged longer than lights. *Percept. Mot. Skills* 39, 63–82. doi: 10.2466/pms.1974.39.1.63
- Grill-Spector, K., Henson, R., and Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends Cogn. Sci.* 10, 14–23. doi: 10.1016/j.tics.2005.11.006

- Grill-Spector, K., Kushnir, T., Edelman, S., Avidan, G., Itzhak, Y., and Malach, R. (1999). Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron* 24, 187–203. doi: 10.1016/S0896-6273(00)80832-80836
- Grill-Spector, K., and Malach, R. (2001). fMRI-Adaptation: a tool for studying the functional properties of human cortical neurons. *Acta Psychol.* 107, 293–321. doi: 10.1016/S0001-6918(01)00019-11
- Hofstötter, C., Mintz, M., and Verschure, P. F. (2002). The cerebellum in action: a simulation and robotics study. *Eur. J. Neurosci.* 16, 1361–1376. doi: 10.1046/j.1460-9568.2002.02182.x
- Ivry, R. B., and Schlerf, J. E. (2008). Dedicated and intrinsic models of time perception. *Trends Cogn. Neurosci.* 12, 273–280. doi: 10.1016/j.tics.2008.04.002
- Kanwisher, N., McDermott, J., and Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J. Neurosci.* 17, 4302–4311.
- Karmarkar, U. R., and Buonomano, D. V. (2007). Timing in the absence of clocks: encoding time in neural network states. *Neuron* 53, 427–438. doi: 10.1016/j.neuron.2007.01.006
- Krekelberg, B., Boynton, G. M., and van Wezel, R. G. A. (2006). Adaptation: from single cells to BOLD signals. *Trends Neurosci.* 29, 250–256. doi: 10.1016/j.tins.2006.02.008
- Lee, K. H., Egleston, P. N., Brown, W. H., Gregory, A. N., Barker, A. T., and Woodruff, P. W. R. (2007). The role of the cerebellum in subsecond time perception: evidence from repetitive transcranial magnetic stimulation. *J. Cogn. Neurosci.* 19, 147–157. doi: 10.1162/jocn.2007.19.1.147
- Leon, M. I., and Shadlen, M. N. (2003). Representation of time by neurons in the posterior parietal cortex of the macaque. *Neuron* 38, 317–327. doi: 10.1016/S0896-6273(03)00185-185
- Lhamon, W. T., and Goldstone, S. (1974). Studies of auditory-visual differences in human time judgment: II. More transmitted information with sounds than lights. *Percept. Mot. Skills* 39, 295–307. doi: 10.2466/pms.1974.39.1.295
- Loffler, G., Yourganov, G., Wilkinson, F., and Wilson, H. R. (2005). fMRI evidence for the neural representation of faces. *Nat. Neurosci.* 8, 1386–1391. doi: 10.1038/nn1538
- Macar, F., Anton, J. L., Bonnet, M., and Vidal, F. (2004). Timing functions of the supplementary motor area: an event-related fMRI study. *Cogn. Brain Res.* 21, 206–215. doi: 10.1016/j.cogbrainres.2004.01.005
- Matell, M. S., and Meck, W. H. (2000). Neuropsychological mechanisms of interval timing behavior. *Bioessays* 22, 94–103. doi: 10.1002/(SICI)1521-1878(200001)22
- Matell, M. S., and Meck, W. H. (2004). Cortico-striatal circuits and interval timing: coincidence detection of oscillatory processes. *Cogn. Brain Res.* 21, 139–170. doi: 10.1016/j.cogbrainres.2004.06.012
- Mauk, M. D., and Buonomano, D. V. (2004). The neural basis of temporal processing. *Annu. Rev. Neurosci.* 27, 307–340. doi: 10.1146/annurev.neuro.27.070203.144247
- Meck, W. H. (1996). Neuropsychology of timing and time perception. *Cogn. Brain Res.* 3, 227–242. doi: 10.1016/0926-6410(96)00009-2
- Meck, W. H. (2005). Neuropsychology of timing and time perception. *Brain Cogn.* 58, 227–242. doi: 10.1016/j.bandc.2004.09.004
- N'Diaye, K., Ragot, R., Garnero, L., and Pouthas, V. (2004). What is common to brain activity evoked by the perception of visual and auditory filled durations? A study with MEG and EEG co-recordings. *Cogn. Brain Res.* 21, 250–268. doi: 10.1016/j.cogbrainres.2004.04.006
- Pariyadath, V., and Eagleman, D. (2007). The effect of predictability on subjective duration. Public Library of Science. *PLoS ONE* 2:e1264. doi: 10.1371/journal.pone.0001264
- Praamstra, P., Kourtis, D., Kwok, H. F., and Oostenveld, R. (2006). Neurophysiology of implicit timing in serial choice reaction-time performance. *J. Neurosci.* 26, 5448–5455. doi: 10.1523/JNEUROSCI.0440-06.2006
- Rao, S. M., Harrington, D. L., Haaland, K. Y., Bobholz, J. A., Cox, R. W., and Binder, J. R. (1997). Distributed neural systems underlying the timing of movements. *J. Neurosci.* 17, 5528–5535.
- Rao, S. M., Mayer, A. R., and Harrington, D. L. (2001). The evolution of brain activation during temporal processing. *Nat. Neurosci.* 4, 317–323. doi: 10.1038/85191
- Roberts, S. (1982). Cross-modal use of an internal clock. *J. Exp. Psychol. Anim. Behav. Process.* 8, 2–22. doi: 10.1037/0097-7403.8.1.2
- Rotshtein, P., Henson, R. N., Treves, A., Driver, J., and Dolan, R. J. (2005). Morphing marilyn into maggie dissociates physical and identity face representations in the brain. *Nat. Neurosci.* 8, 107–113. doi: 10.1038/nn1370
- Siddle, D. A. T. (1991). Orienting, habituation, and resource allocation: an associative analysis. *Psychophysiology* 28, 245–259. doi: 10.1111/j.1469-8986.1991.tb02190.x
- Sokolov, E. N. (1990). The orienting response, and future directions of its development. *Pavlov. J. Biol. Sci.* 25, 142–150. doi: 10.1007/BF02974268
- Teki, S., Grube, M., and Griffiths, T. D. (2011). A unified model of time perception accounts for duration-based and beat-based timing mechanisms. *Front. Integr. Neurosci.* 5:90. doi: 10.3389/fnint.2011.00090
- Treisman, M. (1963). Temporal discrimination and the indifference interval: implications for a model of the internal clock. *Psychol. Monogr.* 77, 1–31. doi: 10.1037/h0093864
- Warm, J. S., Stutz, R. M., and Vassolo, P. A. (1975). Intermodal transfer in temporal Discrimination. *Percept. Psychophys.* 18, 281–286. doi: 10.3758/BF03199375
- Zakay, D. (1989). “Subjective Time and Attentional Resource Allocation: An Integrated Model of Time Estimation,” in *Time and Human Cognition: A life-Span Perspective*, eds I. Levin, and D. Zakay (Oxford: North-Holland), 365–397.
- Zakay, D., Bibi, B., and Algom, D. (2014). Garner interference and temporal information processing. *Acta Psychol.* 147, 143–146. doi: 10.1016/j.actpsy.2013.07.019
- Zakay, D., and Block, R. A. (1996). The role of attention in time estimation processes. *Adv. Psychol.* 115, 143–164. doi: 10.1016/S0166-4115(96)80057-4

**Conflict of Interest Statement:** 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.

Copyright © 2015 Golan and Zakay. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Prior task experience affects temporal prediction and estimation

Simon Tobin\* and Simon Grondin

École de Psychologie, Université Laval, Québec, QC, Canada

## OPEN ACCESS

### Edited by:

Marc Wittmann,  
Institute for Frontier Areas  
of Psychology and Mental Health,  
Germany

### Reviewed by:

Michael Roy,  
Elizabethtown College, USA  
Anne-Claire Rattat,  
Université Fédérale de Toulouse  
Midi-Pyrénées—Centre Universitaire  
Jean-François Champollion, France

### \*Correspondence:

Simon Tobin,  
École de Psychologie, Université  
Laval, Pavillon Félix-Antoine-Savard,  
2325, Rue des Bibliothèques,  
Québec, QC G1V 0A6, Canada  
simon.tobin.1@ulaval.ca,  
simon.grondin@psy.ulaval.ca

### Specialty section:

This article was submitted to  
Perception Science,  
a section of the journal  
Frontiers in Psychology

**Received:** 28 April 2015

**Accepted:** 18 June 2015

**Published:** 06 July 2015

### Citation:

Tobin S and Grondin S (2015)  
Prior task experience affects temporal  
prediction and estimation.  
Front. Psychol. 6:916.  
doi: 10.3389/fpsyg.2015.00916

It has been shown that prior experience with a task improves temporal prediction, even when the amount of prior experience with the task is often limited. The present study targeted the role of *extensive* training on temporal prediction. Expert and intermediate runners had to predict the time of a 5 km running competition. Furthermore, after the race's completion, participants had to estimate their running time so that it could be compared with the predicted time. Results show that expert runners were more accurate than intermediate runners for both predicting and estimating their running time. Furthermore, only expert runners had an estimation that was more accurate than their initial prediction. The results confirm the role of prior task experience in both temporal prediction and estimation.

**Keywords:** timing and time perception, task experience, expert performance, estimation, prediction, running

## Introduction

Time perception, as opposed to other sensory modalities, does not rely on sensory receptors. As a consequence, researchers trying to explain time perception quickly turned into the direction of cognitive processes such as attention and memory (Roetzheim, 2008). While the role of attention in timing has been thoroughly discussed (see Brown, 2008, for a review), some aspects of the involvement of memory, especially long-term memory (LTM), are still understudied, as pointed out recently by many authors (Rattat and Droit-Volet, 2005b; Taatgen and van Rijn, 2011; Tobin and Grondin, 2012). Nonetheless, it should be noted that some aspects of LTM were studied in a timing research perspective, such as the lifespan of time intervals in memory (Gamache and Grondin, 2010), the interference between different temporal traces (Grondin, 2005) or between other task demands and memory traces (Ogden et al., 2008), the development of temporal memory (Rattat and Droit-Volet, 2005a,b, 2007), the effect of the number of presentations of a standard duration on temporal discrimination (Jones and Wearden, 2003; Grondin and McAuley, 2009; Grondin, 2012), the influence of pharmacological substances on temporal memory (Meck, 1983), and the EEG basis of memory traces (Ng et al., 2011).

Even if the involvement of LTM in timing did receive some attention lately, the actual corpus of knowledge in the literature is still thinner than one may wish. One particular overlooked aspect of LTM that has been recently brought up by Tobin and Grondin (2012) is the effect of prior experience with a task on the perceived duration of that task. Indeed, as many daily activities (for instance, driving to work) occurs routinely, it is very likely that one learns temporal information about recurring tasks, temporal information that in turn can improve temporal estimation. As a matter of fact, children as young as 4 years old can classify orderly activities like eating a cookie and watching a movie on the basis of their duration. This shows that children already have a representation of how long some tasks may last (Friedman, 1990).

One of the reason why the influence of prior experience with a task on timing has been overlooked until recently might simply be because it appears too obvious (Tobin and Grondin, 2012). It is

logical to think that one uses experience about a task when such experience is available. Nonetheless, the influence of prior experience on timing clearly deserves empirical investigations for two main reasons. First, as many daily tasks happen more than once, many temporal judgments should occur in situations when prior experience with a task is available. Not taking prior experience with a task into account does not seem a very ecological way to address temporal perception, especially now that a growing number of researchers agree that time perception researches should turn to more ecological tasks (Tobin et al., 2010; Bisson et al., 2012; Matthews and Meck, 2014). Secondly, studying prior experience, as it was shown recently by Tobin and Grondin (2012), sheds light on the involvement of LTM in timing, an involvement that has long been overshadowed by the more prominent and studied role of attention.

### Prior Experience with a Task

The effect of prior experience with a task on timing may be explained by two main cognitive processes. First, as the task is repeated, its execution becomes automatized and requires less attention to perform, leaving more attentional resources for time monitoring. Since the amount of attention available for timing is strongly related to the accuracy of temporal judgments, it explains why the durations of trained tasks are more accurate than novel ones. This demonstration has been reported numerous times in the literature (see Block et al., 2010). The second aspect that could explain the effect of prior experience regards LTM. Indeed, through numerous repetitions of the task, one gains certain knowledge of how long the task lasts.

A recent study by Tobin and Grondin (2012) targeted the involvement of LTM by measuring how different levels of task duration knowledge affect temporal perception. They defined “task duration knowledge” as LTM stored knowledge about the duration of a task. Their study showed that task duration knowledge can improve temporal performance across different temporal tasks (verbal estimation and production) and duration range (from 30 to 90 s). Furthermore, this result was obtained by two distinct manipulations, both requiring the participation of elite athletes (swimmers). First, they compared the temporal perception of two automatized tasks, i.e., tasks with higher task duration knowledge than the other. Secondly, they altered the context in which a single task was performed in order to control the usage of task duration knowledge. In both cases, having more task duration knowledge, or performing the task in a context that allowed relying on task duration knowledge, enhanced the temporal judgments’ precision. In addition, they also performed a third experiment in which elite swimmers were asked to produce 36 s of visualization of a well known task (swimming) and another unknown task (climbing Mount Everest). This experiment further showed that the physical execution is not required to observe an effect of prior experience with a task as the temporal productions of the swimming task (familiar) were much more precise than that of the climbing one (unfamiliar).

While the task was not physically executed in this last experiment, it was still visualized. If no execution at all (whether physically or mentally) is performed, can prior experience with a task still enhance temporal perception? In other words, do elite

athletes like those who participated in Tobin and Grondin (2012) simply know how long it takes them to cover certain distances? The best way to answer that question is to require a temporal prediction of participants with various expertise levels. Indeed, in the prediction task, the temporal judgment is made *before* the task is even executed, thus, the temporal judgment cannot rely on any cues related to the execution of the task but only on previous knowledge with the task at hand. Indeed, the attentional explanation of the effect of prior experience cannot apply to the prediction task; the temporal judgment can only rely on previously learned knowledge stored in LTM.

Thus, the first goal of the experiment is to extend the findings of the Tobin and Grondin (2012) study to the prediction task. In that regards, the literature already provides certain answers. Indeed, many experiments, although they did not use the terms task duration knowledge, did observe the effect of prior experience with a task on temporal prediction (Thomas et al., 2003, 2007; Thomas and Handley, 2008; see Halkjelsvik and Jørgensen, 2012, for a review).

For instance, Thomas et al. (2003) gave participants a little practice time (2 min) with the task before asking them a temporal prediction. It turned out that this simple 2 min of practice strongly increased the prediction performance. Furthermore, Roy et al. (2008) gave participants a single practice trial and further gave a temporal feedback about the duration of that trial for only half of the participants. When asked to make temporal prediction in the following trial, participants who received the temporal feedback were more accurate, showing that they used the information provided by the feedback to guide their next prediction. Finally, Roy and Christenfeld (2007, Experiment 2) compared the prediction of a task based on experience with the task. Indeed, participants had a practice block containing one, three or nine trials of the targeted task (origami). It turned out that the number of trials significantly affected temporal prediction. The number of practice trials affected the side of the error; participants with one practice trial overestimated the time it would take and participants with nine trials underestimated the time it would take.

The aforementioned studies suggest that prior experience with a task increases the precision of the temporal prediction, or changes the side of the error (from overestimation to underestimation). However, it should be noted that, in these experiments, the prior experience is often limited (from only a part of the task to nine repetitions of the task). Although their results were quite interesting, it appears necessary to study the effect of a more *extensive* prior experience. Indeed, as Tobin and Grondin (2012) pointed out in introducing the notion of task duration knowledge, this aspect of temporal perception is relevant for recurring tasks, tasks that are executed on a daily basis, again and again. Thus, although the previously cited experiments were well constructed and have a clear theoretical output, it does not show how the temporal prediction is affected by a level of prior experience that is *comparable* with other daily activities, like driving to work each day for many years.

As a result, one legitimate question arising is the following: what happens when one has an *extensive* training with the task, such as athletes do with their sport? Does the prediction reach an impressive accuracy level, as it is observed with temporal

estimation (see Tobin and Grondin, 2012)? As far as we know, the only study that required the participation of experts (pianists) is the one reported by Boltz et al. (1998; Experiment 2). In their experiment, they compared the time prediction across novices and expert pianists for the execution of musical pieces varying in their degree of familiarity (i.e., identified as recently learned, well learned or extremely well learned). Their results show that for both experts and novices, the degree of familiarity had a significant effect on predicted time: the less familiar the melody was, the longer the predicted duration was. However, experts were surprisingly not better at predicting time than novices, which contradicts what may be expected on the basis of other previously cited studies (Thomas et al., 2003; Roy et al., 2008). Indeed, as prior task experience seems to increase time prediction accuracy, experts should have been better than novices. Two methodological aspects of their experiment may explain this non-significant result. First, participants were instructed to predict to the nearest 30 s. It might have leveled the predictions across the two groups and masked any significant difference that was within a 30-s margin. Furthermore, in music, the key temporal element might be the inter-note interval or tempo, not the overall duration. Hence, it might be best to study the effect of expertise on time prediction with a task in which the elapsed duration is fundamental, like in sports. This idea will be tested in the present experiment.

## Temporal Estimation

Using the prediction task opens the door to studying another relevant aspect of timing. Indeed, while a temporal prediction on its own is interesting, it is even more useful if it is compared with an assessment of the duration once the task is completed. Indeed, as the prediction cannot rely on active time monitoring, it is intriguing to assess how far the prediction is from the temporal estimation of the same task upon completion. Few studies compared directly temporal prediction and the subsequent estimation of the task once completed. Some studies did offer that comparison (Roy and Christenfeld, 2008), but used the retrospective paradigm. Such a paradigm means that participants were not told before the start of the duration to be timed that time estimation would be required. Hence, in a retrospective timing task, participants learn the time estimation requirement afterward. Though retrospective estimates are valuable measures of timing and deserve more empirical investigation (see Tobin et al., 2010; Bisson et al., 2012), it would probably be more relevant, when comparing temporal prediction and estimation, to use the prospective paradigm. In this paradigm, participants are told in advance that a temporal judgment will be required after completing the task. Hence participants can allow more attentional resources to time monitoring, explaining why time estimates in the prospective paradigm are most often reported as more precise than time estimates in retrospective conditions (Block and Zakay, 1997; Block et al., 2010). By using the prospective paradigm, this experiment should answer the following question: if one puts all its attentional resources into timing its running performance, can its estimation be more precise than the initial prediction or there is no gain to be expected?

The few studies left that compared time estimation (prospectively) and prediction do not allow for a clear picture of how these two judgment types differ. First, the Boltz et al.'s (1998) experiment showed that, for expert pianists, the estimated duration was more accurate than the predicted duration. However, the difference between temporal estimation and prediction of novices was mediated by the familiarity with the melody. Indeed, the estimations were more accurate than the predictions for only two of the three familiarity levels (novel and well trained). This improvement was not recorded for the extremely well trained melodies.

On the opposite, Burt and Kemp (1994) found large differences when comparing the prediction and estimation of daily activities (like buying stamps or sorting cards). Indeed, the temporal estimation accuracy after the task's completion was steeply increased when compared to the actual prediction. Hence, the difference between temporal prediction and estimation appears unclear so far and might be mediated by the level of familiarity or prior experience with the task, as suggested by the results of Boltz et al. (1998).

## The Present Study

For the experiment's purpose, expert and intermediate runners were recruited and had to predict how long it would take them to run a 5 km race. Participants were also required to estimate their completion time immediately after the finish line.

Since prior task experience seems to improve temporal judgments, it was expected the more experienced runners to have the best temporal prediction and estimation. Furthermore, we expected the temporal estimation to be more accurate than the initial time prediction as the estimation, being made once the task is completed, could be based on more information (i.e., on how the participants felt, its rank, the fatigue level, etc.).

A third explanatory goal was to assess if all sorts of temporal knowledge are equal. Indeed, runners probably build the task duration knowledge from the feedback they get after each training session (e.g., this session took 43 min). Thus, participants had to report what the sort of feedback they were using (1- measure of time, 2- measure of distance, 3- measure of speed), when training, to see if one sort of feedback provides a better knowledge of one's running time that can translate into more accurate temporal prediction and estimation. We expected that using feedback about speed would be the most efficient feedback type (highest correlations with temporal precision) because one's running speed can be applied to other (i.e., shorter or longer) running situations (e.g., if one knows s/he runs at 10 km/hr, s/he can expect to run 15 km in 90 min).

## Materials and Methods

### Participants

Ninety-one participants (50 males and 41 females) out of the 244 that were registered in a running competition enrolled in the experiment. Six participants were rejected as they did not fill the form properly or did not complete the event, leaving a total of 85 participants. The age of participants ranged from 18 to 66, with a median of 28 years old.

## Material

The participants had to fill out an in-house questionnaire assessing their sporting level, training habits and knowledge of time while running. The questionnaire was in paper form. Three questions measure training habits and were: How often do you get measures of (1) time (2) distance (3) speed when you train? The response scale extended from 1 to 5; 1 = never, 5 = always. They were also asked (on a 1–5 scale, 5 = very well) how well they know the time it takes them to run a specific distance (5 km). The other questions were “You have been participating in running race for how many years?”, “How many times per week do you run?”, “How many hours and minutes per week do you run?”, “What is your running level (amateur, provincial or national)?”, “How many times have you participated in this specific race?” “How far from your real performance would a satisfactory prediction be?”. The runners supplied their own clothing and accessories.

## Procedure

The participants first had to register for the race. The event was a local, on-campus, 5-km race open to the public, although it was also part of a provincial competitive schedule. The circuit consisted of two 2.5-km laps without any distance markers. The circuit changes every year and is unannounced, which means that runners cannot train for this specific race. The goal of the race was to finish not only in the fastest possible time, but in the most accurately predicted time (awards were also given for the best predictions). However, running as fast as possible is still the main goal of the race; the prediction process is simply added for fun. Hence, runners were not simply self-pacing to achieve a good prediction; they ran as fast as they could and hoped they predicted a precise duration. Watches or any other timing devices were prohibited. Each runner stated their predicted time when they registered for the race (and these predictions were later retrieved by the experimenters). After registration, participants were invited to enroll in the experiment. If they accepted, they had to fill out the questionnaire and return it before the start of the race. One of the questions was aimed at defining groups for statistical analyses. Thus, they had to report the level at which they compete: national, provincial, and amateur.

The race proceeded without any intervention on the part of the experimenters. They waited for the runners to pass the finish line before collecting the final running time estimates. The runners knew before the start of the race that this time estimation would be required. The runners took from 924 to 1918 s to complete the race, with a mean time of 1348 s (22 min and 28 s). It should be noted that the weather (early spring in a Northern climate) was particularly difficult with an outside temperature around 4°C wind heavy rain<sup>1</sup> and gusty winds. This study was approved by the *Comité d'éthique de la recherche avec des êtres humains de l'Université Laval*, with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki.

<sup>1</sup>According to verbal reports of many participants, the climate conditions slowed the overall running performance. However, they did report taking the weather into account when registering the temporal prediction.

## Data Analysis

For the purpose of comparing the effect of expertise, two groups were created: experts and intermediates. The expert group consisted of runners who compete regularly at a provincial and national level ( $n = 30$ ). The intermediate group consisted of runners who only compete in amateur events ( $n = 55$ ). This group allocation was based on self-reported information. Hence, in order to investigate if both groups differed in terms of running experience, the amount of training was compared. Expert runners trained in average 4.81 times a week for a total of 6.86 h per week, while these numbers are 3.14 and 2.82, respectively, for intermediate runners. The groups differ significantly for both the number of training sessions per week,  $t(83) = 5.588$ ,  $p < 0.001$ , and the number of training hours per week,  $t(83) = 8.047$ ,  $p < 0.001$ . Furthermore, expert runners reported they have been participating in running races for an average of 8.06 years, while this number goes down to 2.66 for novices. This difference is significant  $t(43.87) = -3.604$ ,  $p < 0.001$ .

Finally, participants were asked to report how well they know the time it takes them to run a specific distance (like 5 km). Expert runners significantly reported a better knowledge ( $M = 4.14$ ) than intermediate runners ( $M = 3.25$ ), with scores on a 1–5 scale (5 = very high). This difference is statistically significant,  $t(78) = 5.106$ ,  $p < 0.001$ . Hence, the distinction between both groups appears adequate since they significantly differ on many aspects<sup>2</sup>.

Two dependent variables were used for assessing performances. The first was the perceived to real time ratio (Ratio), a variable showing the side of the error (over- or underestimation). A Ratio of 1 means a perfect estimation, while Ratios under and over 1 mean time underestimation and overestimation, respectively. The second variable used was the absolute standardized error (ASE), a measure that is not sensitive to the side of the error, thus a more genuine measure of accuracy. The ASE is calculated on the ratio by taking  $|1 - \text{ratio}|$ .

## Results

Table 1 shows the Ratio and ASE for the two time judgments (prediction and estimation), by expertise (experts vs. intermediates). To compare these judgments and assess if the expertise produced an effect of these judgments, a  $2 \times 2$  factorial design ANOVA was first conducted on the Ratio, with time judgment being a repeated-measure factor and expertise a between-subject factor. The ANOVA revealed a significant expertise effect,  $F(1,69) = 7.67$ ,  $p = 0.007$ ,  $\eta^2 = 0.100$  and a significant interaction between time judgment and expertise,  $F(1,66) = 4.55$ ,  $p = 0.036$ ,  $\eta^2 = 0.062$ . A breakdown of the interaction revealed that expert runners were closer to 1 than intermediate for both temporal judgments. Furthermore, for the expert runners, the estimated time was more precise than the predicted time, while there was no difference between these

<sup>2</sup>The experts reported here may not represent “real” experts by some as they are not elite international level. However, the significant differences between the two groups reported here are strong enough to represent two distinct groups having a distinct background with running. It is not a study aimed at extraordinary elite experts.



**TABLE 1 | Mean (*M*) Ratio and ASE as a function of the task and expertise level.**

Expertise level	Task	Ratio		ASE	
		<i>M</i>	SD	<i>M</i>	SD
Intermediates	Prediction	1.039	0.08	0.052	0.05
	Estimation	1.027	0.05	0.043	0.04
Experts	Prediction	0.988	0.02	0.018	0.02
	Estimation	1.004	0.01	0.012	0.01

SD, standard deviation.

two temporal judgments for the intermediate runners. The same ANOVA design was used and conducted on the ASE. This time, only the effect of expertise is significant  $F(1,69) = 13.371$ ,  $p \leq 0.001$ ,  $\eta^2 = 0.109$ , showing that experts are more accurate for both tasks.

Since the previous analyses are based on self reported group attribution, the relation between expertise and temporal performance was further analyzed with correlational analyses. Indeed, correlations between the number of training per week and perceived time were calculated. They show that the more weekly training sessions a runner complete, the more precise the temporal judgments are, and this finding applies to both prediction ( $R = -0.575$ ,  $p \leq 0.001$  for the ASE and  $R = -0.403$ ,  $p \leq 0.001$  for the Ratio) and estimation ( $R = -0.498$ ,  $p \leq 0.001$  for the ASE and  $R = -0.248$ ,  $p = 0.036$  for the Ratio).

Furthermore, runners were asked to report the frequency to which they use measures of distance, time, and speed. Correlational analyses were conducted to assess if the use of a specific feedback was associated with temporal accuracy (again using the percentage of error). The analyses revealed that the use of speed was the only feedback type that correlated significantly with time prediction ( $R = -0.285$ ,  $p = 0.019$  for the ASE and  $R = -0.239$ ,  $p = 0.033$ ). Thus, the more runners reported using measures of running speed while training (regardless of their expertise levels), the more precise was their predicted time. A mediation analysis revealed that the use of speed-related feedbacks did not mediate the effect of expertise. Although correlated to predicted time, the usage of feedback was not correlated to estimated time.

## Discussion

This section will first discuss about the effect of extensive training on temporal performance and secondly, will contrast the prediction and the subsequent estimation.

### Effect of Experience

The results show that expert runners are better at predicting their running time than intermediate runners. This conclusion is coherent with other studies showing prior task experience enhances the prediction accuracy (Boltz et al., 1998). While there is sufficient body of studies showing this role of prior task experience on temporal prediction, the demonstrations were usually made with very limited prior task experience or training with the task. Hence, the participation of experienced runners allowed assessing how extensive training affects the accuracy of the prediction.

Both groups of runners exhibited surprisingly unbiased predictions. Indeed, compared to other studies using temporal prediction (see 1 in Roy et al., 2005), the ratios recorded here are quite close to 1. Hence, it suggests that the more one is experienced with a task, the better the prediction becomes. That is coherent with Tobin and Grondin's (2012) study in which experimented athletes reached an accuracy level on a temporal estimation task much better than what is generally observed in the literature for similar tasks/durations. Consequently, both studies converge and show that temporal perception processes (estimation or prediction) are strongly affected by prior task experience and that a "near-perfect" ratio is possible with sufficient training with the temporal task.

Another aspect of the results is interesting. Indeed, not only were the expert runners more accurate, the side of their error (over- or underestimation) was the opposite than the one observed with intermediate runners. Indeed, expert runners predicted a faster performance than what they actually accomplished while intermediate runners underestimated their performance by predicting a slower time. The amount of prior task experience not only affected temporal precision, but also caused a directional effect. This directional effect may be caused by one's confidence into personal abilities, with experts being more confident than intermediates.

While these results show that experts are better at perceiving time, little is known as to *why* exactly they are better. Tobin et al. (2010) studied the time perception of gamers for 12 and 35 min of gaming. In their studies, gamers reported playing an amount of 12.95 h per week on average. This amount of game play exceeds by far the amount of training reported here by the runners. However, gamers were quite imprecise at estimating time, with ratios ranging from 1.2 to 1.6 depending on the duration used. Thus, it appears that doing a specific task often, be it running or playing video games, is not sufficient to create temporal expertise. The main difference between these two tasks can be the importance of time. Indeed, when playing, the duration of the game is not important. In fact, many players reported they specifically play to lose track of time (Wood et al., 2007). However, when training, runners may pay close attention to their distance, time and speed. Hence, for the large experience with the task to translate into more accurate temporal perception, it might be necessary to pay attention to the duration of each activity (i.e., each training session) and get timely feedback (e.g., this 5-km training took 21 min). Without these feedbacks, temporal expertise may not develop, like in the case of gamers. Indeed, it is well known that temporal feedback improves time perception (Fraisse, 1971; Hicks and Miller, 1976; Ryan and Fritz, 2007).

This explication is also coherent with the memory bias account proposed by Roy et al. (2005). Indeed, they suggested that the inaccuracy in the temporal prediction could be caused by an inaccuracy in the memory of the previous occurrence of the task. Said differently, people have poor prediction because they remembered poorly the duration it took the previous times. Thus, receiving timely feedback may often help creating an accurate memory of how long the task last, which in turn translates into accurate predictions. In line with this idea, runners had to report what kind of feedbacks, if any, they use while training (elapsed

duration, traveled distance, or averaged speed) to see if the use of these feedbacks correlates with temporal performance.

The results show that, among time, speed and distance, it is only the usage of speed-related feedback that is significantly correlated with the accuracy of time prediction, regardless of the expertise level. Hence, the more a runner uses a measure of speed when training, the more precise at predicting time he/she becomes. This finding suggests that runners could gain their temporal expertise through the feedback they got after each training session (in fact, many GPS systems nowadays seem to have this in mind, helping runners know their running pace when training). Indeed, by learning their average speed, it becomes easier for them to know how long running a specific distance should take by using a simple formula based on their average speed.

### Prediction vs. Estimation

The second main goal of the experiment was to contrast the initial temporal prediction with the estimation upon completion. As stated in the introduction, few studies compared the performance level on a temporal prediction with its subsequent temporal estimation. Furthermore, the conclusions from such studies differed, offering quite a complex picture. Based on our results, both the accuracy of the initial prediction and the expertise level of the participant might explain the difference between prediction and estimation and further explain why different studies reached different conclusions.

First, for novel or occasional tasks such as the one used by Burt and Kemp (1994), the recorded predictions were far from accurate. Hence, once the task is completed, participants may easily figure that their prediction was wrong and adjust it with a more precise estimation. This could explain why in such cases the estimation is more accurate than the prediction. Indeed, the farther the prediction is from the actual duration, the larger are the chances to improve the subsequent temporal estimation as there is much more room for improvement.

However, when the prediction accuracy is closer to the target duration, it may take a certain level of expertise to be able to adjust that prediction and make a more precise estimation. Indeed, our intermediate runners did not improve their prediction accuracy when estimating time after completion. Similarly, novice pianists in the Boltz et al.'s (1998) experiment only improved in the 2° of familiarity (novel and well trained) for which their predicted time was the less accurate (however, for the extremely well trained melody, the prediction of novices was more accurate and their estimation did not improve that prediction). On the opposite, our expert runners and Boltz et al.'s (1998) expert pianists were always better at estimating than predicting time, even if they were better than novices at predicting time. Hence, it may require a certain level of expertise with the task in order to “read” the duration of the task and correct the prediction into a more precise temporal

estimation. Thus not only prior task experience or expertise would predict the accuracy of the prediction, but it would also predict one's ability to make a temporal estimation that is more accurate than its initial prediction.

### Limitations and Future Studies

Relying on athletes allowed testing an amount of training that is almost impossible to recreate in a laboratory setting. As the insufficient amount of training in other studies to fully reflect “real-life” situations was an important issue, the participation of athletes was a sound choice. However, the clear drawback from this decision is that participant came to the study with their own background; it was thus impossible to monitor their training. Since we advocate for more ecological studies in timing (see Tobin et al., 2010), especially when studying prior task experience, we argue that this limitation is minor. However, subsequent studies with more experimental control on the training process will be necessary to better understand how prior experience improves timing. Especially, monitoring the training process could be very informative and might show the learning curve (for instance, what amount of training is required to reach an asymptotic temporal performance?).

It could be argued that another limitation of the present study is the fact that the groups were separated on the basis of self-reported data (expert or intermediate). However, the statistical analysis made on the amount of training actually showed both groups do differ significantly. Furthermore, correlation analyses showed that the more runners train, the more accurate their temporal perception is. This key finding is independent from the group attribution.

### Conclusion

This study adds to the large body of evidence showing that prior task experience enhances temporal prediction accuracy. Furthermore, the participation of athletes showed that with more experience with a task, predictions get more accurate. It further shows that extensive training improves temporal performance up to an impressive level. This finding also applies to the temporal estimation made after the task's completion. Finally, the difference between the prediction and the estimation of a task may depend on both the accuracy of the prediction, and the level of experience with a task.

### Acknowledgments

This project was funded by a grant awarded to ST by the National Science and Engineering Research Council of Canada. The authors would like to thank Richard Chouinard for his collaboration with the project.

### References

- Bisson, N., Tobin, S., and Grondin, S. (2012). Prospective and retrospective time estimates of children: a comparison based on ecological tasks. *PLoS ONE* 7:e33049. doi: 10.1371/journal.pone.0033049
- Block, R. A., Hancock, P. A., and Zakay, D. (2010). How cognitive load affects duration judgments: a meta-analytic review. *Acta Psychol.* 134, 330–343. doi: 10.1016/j.actpsy.2010.03.006
- Block, R. A., and Zakay, D. (1997). Prospective and retrospective duration judgments: a meta-analytic review. *Psychon. Bull. Rev.* 4, 184–197. doi: 10.3758/BF03209393

- Boltz, M. G., Kupperman, C., and Dunne, J. (1998). The role of learning in remembered duration. *Mem. Cogn.* 26, 903–921. doi: 10.3758/BF03201172
- Brown, S. W. (2008). “Time and attention: review of the literature,” in *Psychology of Time*, ed. S. Grondin (Bingley: Emerald Group Publishing), 111–138.
- Burt, C. D. B., and Kemp, S. (1994). Construction of activity duration and time management potential. *Appl. Cogn. Psychol.* 8, 155–168. doi: 10.1002/acp.2350080206
- Fraisse, P. (1971). L'apprentissage de l'estimation de la durée et ses repères. *Année Psychol.* 71, 371–379. doi: 10.3406/psy.1971.27748
- Friedman, W. J. (1990). Children's representations of pattern of daily activities. *Child Dev.* 61, 1399–1412. doi: 10.2307/1130751
- Gamache, P.-L., and Grondin, S. (2010). The lifespan of time intervals in reference memory. *Perception* 39, 1431–1451. doi: 10.1068/p6652
- Grondin, S. (2005). Overloading temporal memory. *J. Exp. Psychol. Hum. Percept. Perform.* 31, 869–879. doi: 10.1037/0096-1523.31.5.869
- Grondin, S. (2012). Violation of the scalar property for time perception between 1 and 2 seconds: evidence from interval discrimination, reproduction, and categorization. *J. Exp. Psychol. Hum. Percept. Perform.* 38, 880–890. doi: 10.1037/a0027188
- Grondin, S., and McAuley, J. D. (2009). Duration discrimination in crossmodal sequences. *Perception* 38, 1542–1559. doi: 10.1068/p6359
- Halkjelsvik, T., and Jørgensen, M. (2012). From origami to software development: A review of studies on judgment-based predictions of performance time. *Psychol. Bull.* 138, 238–271. doi: 10.1037/a0025996
- Hicks, R. E., and Miller, G. W. (1976). Transfer of time judgments as a function of feedback. *Am. J. Psychol.* 89, 303–310. doi: 10.2307/1421413
- Jones, L. A., and Wearden, J. H. (2003). More is not necessarily better: examining the nature of the temporal reference memory component in timing. *Q. J. Exp. Psychol. B Comp. Physiol. Psychol.* 56B, 321–343. doi: 10.1080/02724990244000287
- Matthews, W. J., and Meck, W. H. (2014). Time perception: the bad news and the good. *WIREs Cogn. Sci.* 5, 429–446. doi: 10.1002/wcs.1298
- Meck, W. H. (1983). Selective adjustment of the speed of internal clock and memory processes. *J. Exp. Psychol. Anim. Behav. Process.* 9, 171–201. doi: 10.1037/0097-7403.9.2.171
- Ng, K. K., Tobin, S., and Penney, T. B. (2011). Temporal accumulation and decision processes in the duration bisection task revealed by contingent negative variation. *Front. Integr. Neurosci.* 5:77. doi: 10.3389/fnint.2011.00077
- Ogden, R. S., Wearden, J. H., and Jones, L. A. (2008). The remembrance of times past: interference in temporal reference memory. *J. Exp. Psychol. Hum. Percept. Perform.* 34, 1524–1544. doi: 10.1037/a0010347
- Rattat, A. C., and Droit-Volet, S. (2005a). La mémoire à long terme des durées: fonctionnement et développement. *Psychol. Fr.* 50, 99–116. doi: 10.1016/j.psfr.2004.10.006
- Rattat, A. C., and Droit-Volet, S. (2005b). The long-term retention of time: a developmental study. *Q. J. Exp. Psychol. B Comp. Physiol. Psychol.* 58B, 163–176. doi: 10.1080/02724990444000096
- Rattat, A. C., and Droit-Volet, S. (2007). Implicit long-term memory for duration in young children. *Eur. J. Cogn. Psychol.* 19, 271–285. doi: 10.1080/09541440600834647
- Roeckelein, S. W. (2008). “Time and attention: review of the literature,” in *Psychology of Time*, ed. S. Grondin (Bingley: Emerald Group Publishing), 111–138.
- Roy, M. M., and Christenfeld, N. J. S. (2007). Bias in memory predicts bias in estimation of future task duration. *Mem. Cogn.* 35, 557–564. doi: 10.3758/BF03193294
- Roy, M. M., and Christenfeld, N. J. S. (2008). Effect of task length on remembered and predicted duration. *Psychon. Bull. Rev.* 15, 202–207. doi: 10.3758/PBR.15.1.202
- Roy, M. M., Christenfeld, N. J. S., and McKenzie, C. R. M. (2005). Underestimating the duration of future events: memory incorrectly used or memory bias? *Psychol. Bull.* 131, 738–756. doi: 10.1037/0033-2909.131.5.738
- Roy, M. M., Mitten, S. T., and Christenfeld, N. J. S. (2008). Correcting memory improves accuracy of predicted task duration. *J. Exp. Psychol. Appl.* 14, 266–275. doi: 10.1037/1076-898X.14.3.266
- Ryan, L. J., and Fritz, M. S. (2007). Erroneous knowledge of results affects decision and memory processes on timing tasks. *J. Exp. Psychol. Hum. Percept. Perform.* 33, 1468–1482. doi: 10.1037/0096-1523.33.6.1468
- Taatgen, N., and van Rijn, H. (2011). Traces of times past: representations of temporal intervals in memory. *Mem. Cogn.* 39, 1546–1560. doi: 10.3758/s13421-011-0113-0
- Thomas, K. E., and Handley, S. J. (2008). Anchoring in time estimation. *Acta Psychol.* 127, 24–29. doi: 10.1016/j.actpsy.2006.12.004
- Thomas, K. E., Handley, S. J., and Newstead, S. E. (2007). The role of prior task experience in temporal misestimation. *Q. J. Exp. Psychol.* 60, 230–240. doi: 10.1080/17470210600785091
- Thomas, K. E., Newstead, S. E., and Handley, S. J. (2003). Exploring the time prediction process: the effects of task experience and complexity on prediction accuracy. *Appl. Cogn. Psychol.* 17, 655–673. doi: 10.1002/acp.893
- Tobin, S., Bisson, N., and Grondin, S. (2010). An ecological approach to prospective and retrospective timing of long durations: a study involving gamers. *PLoS ONE* 5:e9271. doi: 10.1371/journal.pone.0009271
- Tobin, S., and Grondin, S. (2012). Time perception is enhanced by task duration knowledge: evidence from experienced swimmers. *Mem. Cogn.* 40, 1339–1351. doi: 10.3758/s13421-012-0231-3
- Wood, R. T. A., Griffiths, M. D., and Parke, A. (2007). Experiences of time loss among videogame players: an empirical study. *Cyberpsychol. Behav.* 10, 38–44. doi: 10.1089/cpb.2006.9994

**Conflict of Interest Statement:** 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.

Copyright © 2015 Tobin and Grondin. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Crossmodal Statistical Binding of Temporal Information and Stimuli Properties Recalibrates Perception of Visual Apparent Motion

Yi Zhang<sup>1</sup> and Lihan Chen<sup>1,2\*</sup>

<sup>1</sup> Department of Psychology and Beijing Key Laboratory of Behavior and Mental Health, Peking University, Beijing, China,

<sup>2</sup> Key Laboratory of Machine Perception (Ministry of Education), Peking University, Beijing, China

## OPEN ACCESS

### Edited by:

Laurence T. Maloney,  
Stanford University, USA

### Reviewed by:

Sergei Gepshtein,  
Salk Institute for Biological Studies,  
USA

Takahiro Kawabe,  
Nippon Telegraph and Telephone  
Corporation, Japan

### \*Correspondence:

Lihan Chen  
clh@pku.edu.cn

### Specialty section:

This article was submitted to  
Perception Science,  
a section of the journal  
Frontiers in Psychology

**Received:** 16 July 2015

**Accepted:** 11 March 2016

**Published:** 29 March 2016

### Citation:

Zhang Y and Chen L (2016)  
Crossmodal Statistical Binding  
of Temporal Information and Stimuli  
Properties Recalibrates Perception  
of Visual Apparent Motion.  
*Front. Psychol.* 7:434.  
doi: 10.3389/fpsyg.2016.00434

Recent studies of brain plasticity that pertain to time perception have shown that fast training of temporal discrimination in one modality, for example, the auditory modality, can improve performance of temporal discrimination in another modality, such as the visual modality. We here examined whether the perception of visual Ternus motion could be recalibrated through fast crossmodal statistical binding of temporal information and stimuli properties binding. We conducted two experiments, composed of three sessions each: pre-test, learning, and post-test. In both the pre-test and the post-test, participants classified the Ternus display as either “element motion” or “group motion.” For the training session in Experiment 1, we constructed two types of temporal structures, in which two consecutively presented sound beeps were dominantly (80%) flanked by one leading visual Ternus frame and by one lagging visual Ternus frame (VAAV) or dominantly inserted by two Ternus visual frames (AVVA). Participants were required to respond which interval (auditory vs. visual) was longer. In Experiment 2, we presented only a single auditory–visual pair but with similar temporal configurations as in Experiment 1, and asked participants to perform an audio–visual temporal order judgment. The results of these two experiments support that statistical binding of temporal information and stimuli properties can quickly and selectively recalibrate the sensitivity of perceiving visual motion, according to the protocols of the specific bindings.

**Keywords:** Ternus display, temporal structure, intersensory binding, statistical learning, interval

## INTRODUCTION

In a typical temporal ventriloquism effect, perception of the onset of a visual event or the intervals of paired visual events is biased by the presentation of nearby auditory clicks or paired auditory beeps (Chen and Vroomen, 2013). For example, Morein-Zamir et al. (2003) showed that when presenting a sound before the first light and a second sound after the second light (the AVVA configuration), participants could more easily differentiate the two lights, as if the sounds pulled the lights further apart in time. In contrast, when the two sounds occurred in between the two lights, the sounds apparently pulled the lights closer together and made it difficult to judge the order of visual lights, rendering participants’ performance less accurate (Morein-Zamir et al., 2003). The temporal ventriloquism effect has recently been extended to dynamic scenarios by employing the visual Ternus display (Shi et al., 2010). The Ternus display involves a multi-element stimulus



that can induce either of two different percepts of apparent motion: “element motion” or “group motion.” In this study, each frame had two disks, with the second disk of the first frame and the first disk of the second frame being presented at the same location. The perception of “element motion” or “group motion” is dependent on the perceived interval between the two Ternus frames. When the inter-frame interval is short, observers perceive “element motion,” in which the endmost disk is seen as moving back and forth while the middle disk, at the central position, remains stationary or flashing. When the inter-frame interval is longer, observers generally perceive “group motion,” in which both disks appear to move laterally as a whole. The two perceptions are mutually exclusive. The visual Ternus display thus provides a good tool for manipulating crossmodal temporal disparities. This study also found that two sounds presented in temporal proximity to, or synchronously with, the two visual frames, respectively, can shift the transitional threshold for visual apparent motion. However, such effects were not evident with single-sound configurations (Shi et al., 2010).

Temporal perception bias has been demonstrated not just in the one trial demonstration of audiovisual integration, but also after an adaptation procedure. Here, the auditory and visual events each occur separately beyond the time window in which multisensory integration could have taken place (Spence and Squire, 2003). Using a temporal adaptation task and employing the Ternus apparent motion as probes, Zhang et al. (2012) found that adapting to different time intervals conveyed through stimuli in different modalities affects the subsequent implicit perception of visual timing. The stimuli in different modalities could be frames of a visual Ternus display, visual blinking disks, or auditory beeps. Adapting to the short time interval in all of the above situations led to more reports of “group motion” for the subsequent Ternus display. However, adapting to the long time interval gave rise to different results. In this condition, no aftereffects for visual adaptation occurred, while there were significantly more reports of group motion for auditory adaptation (Zhang et al., 2012). Additionally, Chen and Zhou (2014), also using the Ternus apparent motion as probes, examined the extent to which the ability to discriminate sub-second time intervals acquired in one sensory modality can be transferred to another modality with a fast perceptual training protocol. The training protocol required participants to explicitly compare the interval length between a pair of visual, auditory, or tactile stimuli with a standard interval. Results showed that after fast explicit training of interval discrimination (about 15 min), participants improved their ability to categorize the visual apparent motion in the Ternus displays. However, the training benefits here were mild for visual timing. Overall, in light of the evidence of crossmodal transfer of time perception and adaptation, it seems a central clock may account for sub-second temporal processing (Ivry and Schlerf, 2008; Chen and Zhou, 2014).

Beyond temporal manipulations, previous studies have investigated the role of feature binding in crossmodal time perception. Evidence so far supports that a single auditory event can selectively bind with only one of multiple visual events, or alternatively, interact with all of the visual events, to

reach a perceptual decision (simultaneity judgment or feature discrimination) on them (Van der Burg et al., 2008, 2013; Roseboom et al., 2009, 2013). This flexible association of temporal pairings is also shown in a person’s own actions and sensory feedback. In this case, exposing the left and right hands to different action-effect lags can concurrently lead to different amounts of the temporal recalibration effect (Sugano et al., 2014).

The different and selective adaptation reported in the above studies has indeed addressed the aftereffects of fixed temporal relations between different sensory events or between the action and its feedback. The current study asks whether perception of time intervals in one modality can be implicitly biased by inferring temporal relations between crossmodal events, in which the observers should use both the temporal information and stimuli properties. The statistical binding of temporal information and stimuli properties, implemented through presentations of probable audiovisual events, would let the observers form a temporary prior assessment of the temporal (interval) relations between the target events. Hence, the observable temporal aftereffects would be rendered. Moreover, statistical binding of temporal information and stimuli properties could largely form strong temporal perceptual groupings, which would otherwise be less obvious or absent with single or fewer trials of the presentation of audiovisual pairs (see Experiment 2 in Shi et al., 2010). In the present study, we investigated this hypothesis by constructing selective temporal relations between visual Ternus frames (with black or red elements) and auditory beeps. We expected that the temporal interval modulations between the paired auditory beeps and the visual Ternus frames would give rise to different adaptation aftereffects. This would then lead to different biases of perceiving “element motion” vs. “group motion” in the post-test of the Ternus display. We conducted two experiments, detailed below, to examine our hypotheses.

## MATERIALS AND METHODS

The procedure of pre-test, training, and post-test was adopted. The pre-test and post-test tasks were discriminations of visual Ternus apparent motion (“element motion” vs. “group motion”). The interim training sessions were tasks of temporal discrimination of auditory–visual events—either interval comparison (Experiment 1) or temporal order judgment (TOJ; Experiment 2).

## EXPERIMENT 1

In Experiment 1, we manipulated the temporal interval structure between paired auditory–visual events. We set up two configurations of the Ternus display. That is, the Ternus frame contained either two black disks or two red disks. We paired mostly (80% of total trials) the black frames with a temporal structure in which two visual Ternus frames were inserted between two auditory beeps (the VAAV configuration). Meanwhile, two red frames were mainly associated with another

temporal structure—two beeps were inserted between two visual Ternus frames (the AVVA configuration). We hypothesized that the statistically dominant VAAV configuration would lead to a decrease in sensitivity for visual intervals, and this influence would generalize to the Ternus motion [with increased just noticeable differences (JNDs)]. In contrast, the dominant AVVA structure would lead to an increase in sensitivity for visual intervals, decreasing the JNDs for judging Ternus motion in the post-test.

## Participants

Twenty-eight students (15 females) from Peking University took part in Experiment 1. The mean age of the sample was 22.1 years old. Seventeen students (nine females) attended Experiment 1a, in which the sample had a mean age of 21.9 years old. Eleven students (six females) participated in Experiment 1b, in which the sample had a mean age of 22.2 years old. All the participants had normal or corrected-to-normal vision, and normal hearing. All were naïve as to the purpose of the experiment. The study was approved by the Ethics Committee of the Department of Psychology at Peking University and informed consent was obtained before the experiment for all participants.

## Stimuli and Apparatus

The visual stimuli consisted of two frames, each containing two disks ( $1.3^\circ$  of visual angle in diameter) presented on a gray background ( $16.1 \text{ cd/m}^2$  luminance). The disks were either red ( $10.6 \text{ cd/m}^2$  luminance) or black ( $12.7 \text{ cd/m}^2$  luminance) and the disks in each trial were of the same color. The separation between the two disks was  $2^\circ$  of visual angle. As shown in **Figure 1**, the two frames shared one element location at the center of the monitor but contained two other elements located at horizontally opposite positions relative to the center.

Visual stimuli were presented on a 22-inch CRT monitor ( $1,024 \times 768$  pixels; 100 Hz) controlled by a PC (HPAMD Athlon 64 Dual-Core Processor) with a Radeon 1700 FSC graphics card. Viewing distance was set to 57 cm, maintained by using a chinrest. The testing room was dimly lit with an average ambient luminance of  $0.12 \text{ cd/m}^2$ . Audio stimuli (65 dB, 1,000 Hz) were generated and delivered via an M-Audio card (Delta 1010) to a headset (RT-788V, RAPTOXX). Stimulus presentation and data collection were implemented by computer programs which were

developed with Matlab 7.1 (MathWorks Inc., Natick, MA, USA) and Psychophysics Toolbox (Brainard, 1997; Pelli, 1997).

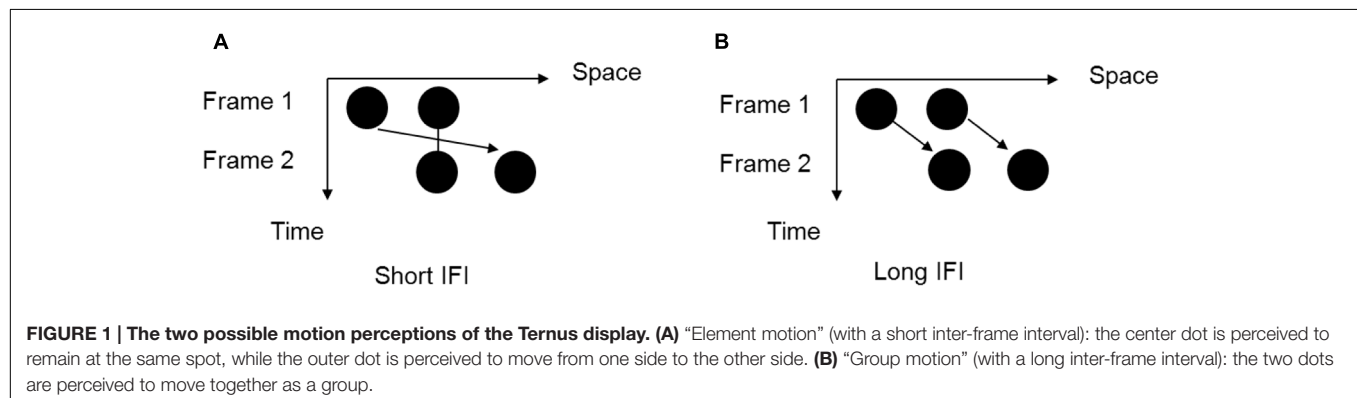
## Design and Procedure

A between-participants design was adopted in Experiment 1. Experiment 1a was composed of three sections: pre-test, training, and post-test. In the pre-test and post-test, a Ternus display showing either two black frames or two red frames was used. In Experiment 1b (the control test for Experiment 1a), the pre-test and post-test were the same as in Experiment 1a, except that the participants were required to take a rest during the time equivalent to that of the training session in Experiment 1a.

### Pre-test and Post-test

Before the formal experiment, participants underwent practice to become familiar with a Ternus display of both the typical element motion (ISI = 50 ms) and group motion (ISI = 200 ms) percepts. They were asked to discriminate the above two percepts by pressing the left and right mouse button to indicate judgments of each element motion and group motion, respectively. The mapping between button and response type was counterbalanced across participants. When participants made an incorrect response, immediate feedback appeared on the screen showing the percept (element motion or group motion) that they should have reported. The practice session continued until the participant's report accuracy was close to 100%. Almost all of the participants met this standard within 120 trials.

In the pre-test phase, each trial began with a fixation cross presented at the center of the display for 300 ms, followed by a blank display with a random interval of 500–700 ms. Next, typical Ternus motion was depicted as two frames with a random ISI (50, 80, 110, 140, 170, 200, or 230 ms) as shown in **Figure 1**. Each Ternus frame was presented for 30 ms. After another blank display for 500 ms, a question mark appeared to prompt participants to make a forced-choice response by using the mouse button. The next trial began 500 ms after the participant pressed the button. There were 24 trials for each ISI level. Color (red or black) and the directions of apparent motion (leftward or rightward) were balanced across trials. The 336 trials were divided into four blocks and participants could take a short rest between blocks. There was no feedback in the pre-test session.

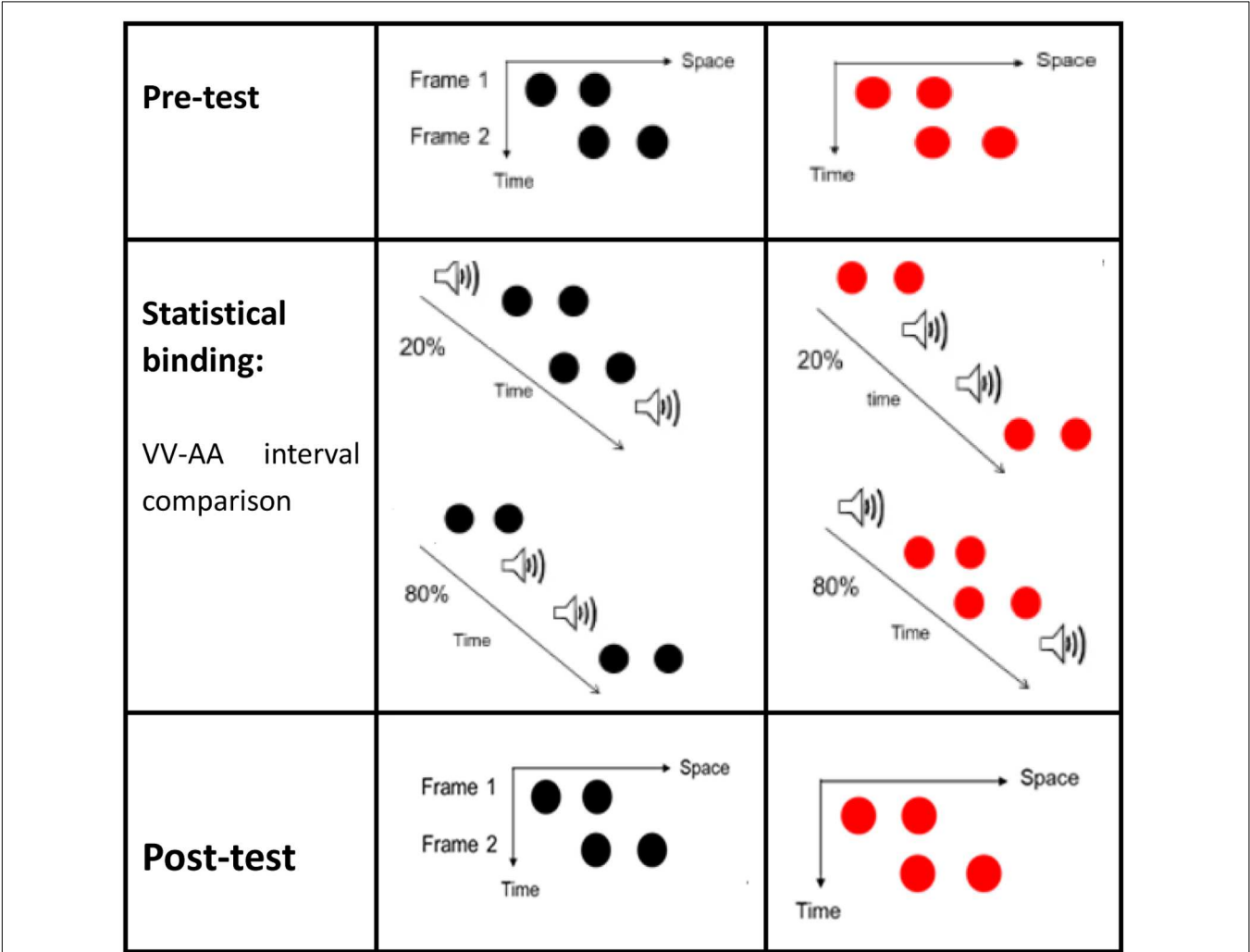


The procedure of the post-test phase was the same as that of the pre-test phase.

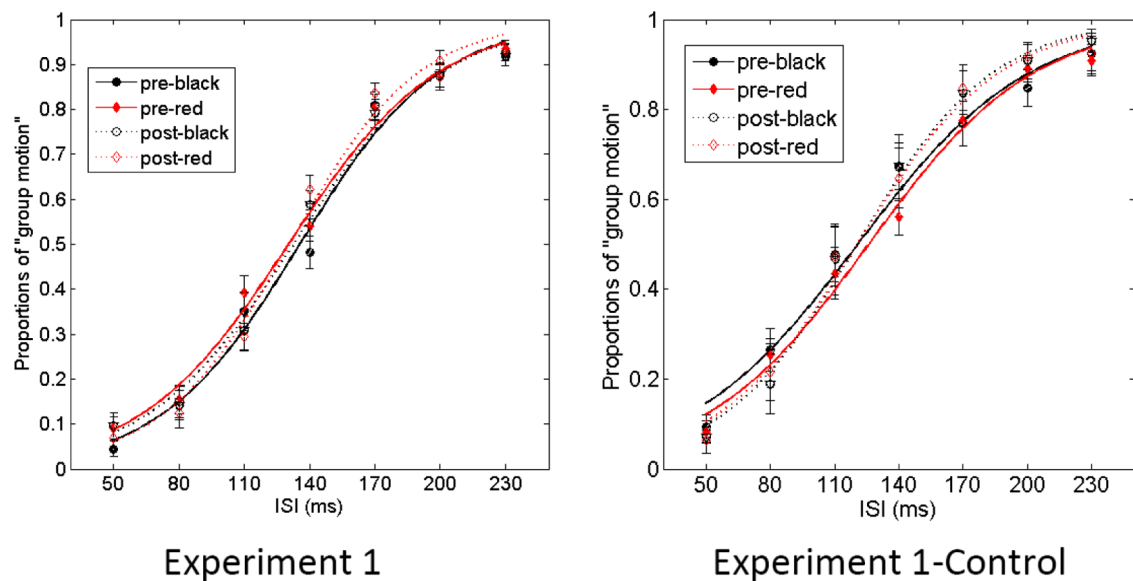
Training

Participants in the training group were required to complete an interim training session on temporal interval discriminations before the post-test. After participants saw a fixation cross for 300 ms and a blank display for 300–500 ms, two frames would appear on the screen with a random ISI of 50–230 ms. Each frame contained two disks presented consecutively at the center of the screen. The color of the disks in a given trial was either black or red. Two brief 30 ms sound beeps appeared along with the two visual stimuli. There were two conditions of the audiovisual interval. In condition 1, for 80% of the trials, the first sound preceded the first red visual frame and the second sound trailed

the second red visual frame by 80 ms (the AVVA condition). **Figure 2** shows that in condition 2, for 80% of the trials, the first sound trailed the first black visual Ternus frame by 80 ms and the second sound preceded the second black visual frame by 80 ms. These were called “inner sounds” (VAAV temporal structure). For the less common condition (20% of trials), the temporal structures of AVVA and VAAV were used in reverse to those in the more common condition (80% trials). When the audiovisual stimuli were presented again, after another blank screen of 500–700 ms, text appeared on the screen asking “Which time interval is longer, the visual or the sound beep?” Participants were then prompted to respond by pressing the associated mouse key (the left key to indicate that the auditory interval was longer and the right key to indicate that the visual interval was longer). When they were incorrect, immediate feedback would appear on the



**FIGURE 2 | Ternus displays (for the pre-test and post-test) and illustrations of the stimuli configurations for Experiment 1.** Two kinds of Ternus displays were used (with black frames or red frames). In the training session, for 80% of the red-red (RR) configuration trials, the first sound preceded the first red visual frame and the second sound trailed the second red visual frame by 80 ms (hereafter referred to simply as the AVVA condition). In 80% of the black-black (BB) configuration trials, the first visual frame preceded the first beep and the second visual frame trailed the second beep by 80 ms (hereafter referred to simply as the VAAV condition). The inter-stimulus-interval (ISI) between the two Ternus frames was randomly set from between 50 and 230 ms inclusive. For the other 20% of the trials, the RR and BB configurations were associated with temporal structures of VAAV and AVVA, respectively.



**FIGURE 3 | Average psychometric curves for Experiment 1.** In the graph “Experiment1” (Left), the solid line with solid black circles shows the percentage of “group motion” for the pre-test of the black Ternus display. The solid line with solid red diamonds indicates the pre-test of the red Ternus display. The dotted line with empty circles indicates the percentage of “group motion” for the post-test of the black Ternus display, and the dotted line with empty diamonds the post-test of the red Ternus display. In the graph “Experiment 1-Control” (Right), the various colored curves indicate the same as in the left graph. The error bars in both graphs represent the SEM.

screen telling them so. The next trial began 1–1.2 s after the participant pressed the button. A total of 360 trials was divided into six blocks between which participants could take a short rest.

## Results

### Pre-test and post-test

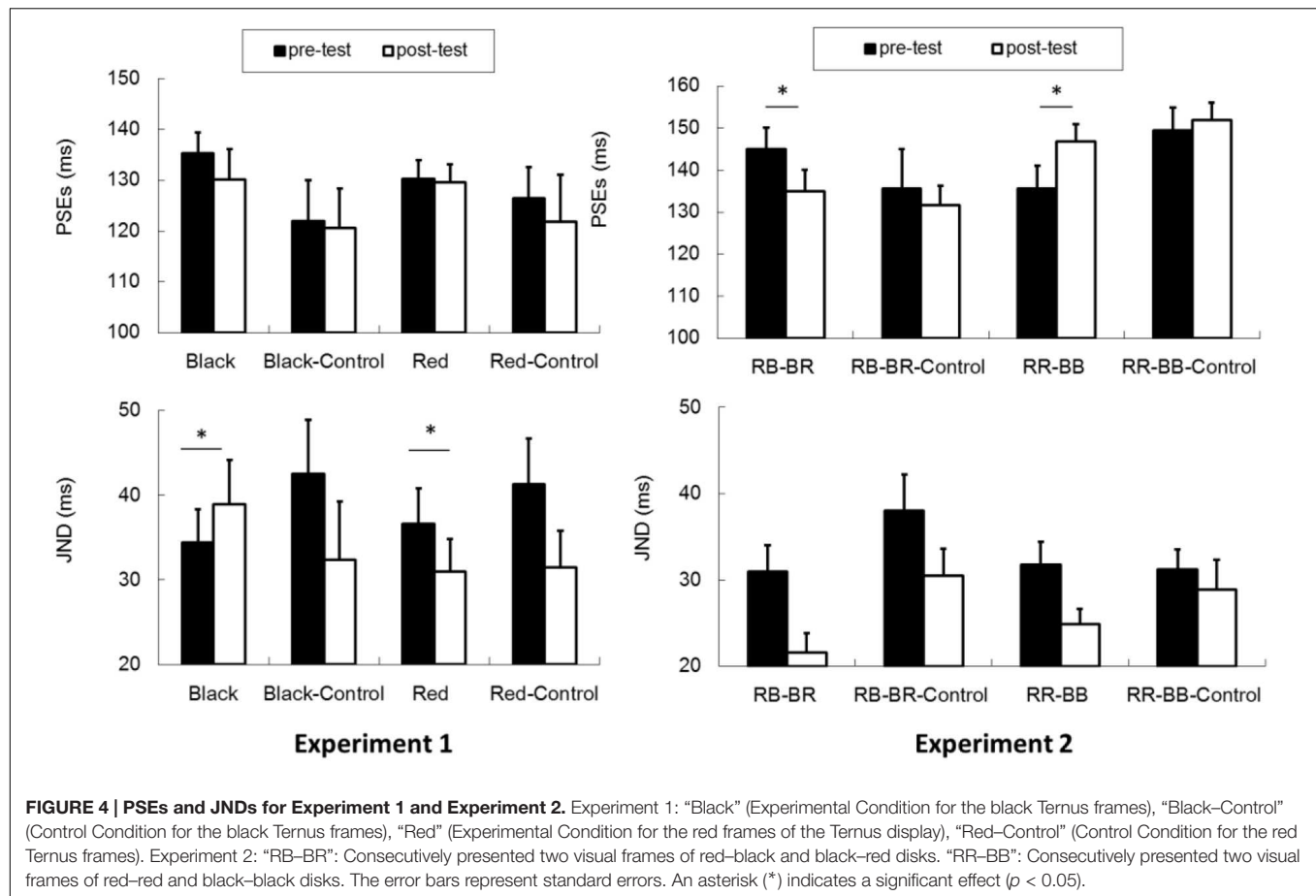
The proportion of group motion reports was plotted as a function of ISI and fit by a logistic regression for each participant (Figure 3). For each condition (black vs. red Ternus frame), the transitional threshold between element motion and group motion, that is, the point at which group motion and element motion were reported with equal frequency, was calculated by estimating the 50% performance point on the (fitted) logistic function. The transitional threshold is also referred to as the point of subjective equality (PSE). The just noticeable difference (JND) represents the difference between the two motion perceptions, which is obtained by estimating the ISI difference of half between 25% and 75% of the group motion responses from the psychometric curves (Treutwein and Strasburger, 1999). PSEs were  $135.3 \pm 4.1$  (SE) and  $130.3 \pm 3.7$  for the pre-test of the black Ternus type and the pre-test of the red Ternus type, respectively. PSEs were  $130.1 \pm 6.1$  and  $129.6 \pm 3.6$  for the post-test of the black Ternus type and the post-test of the red Ternus type, respectively. A repeated measures analysis of variance (ANOVA) with independent factors [Ternus type: Color (black or red) vs. Time (pre-test or post-test)] showed that the PSEs were statistically equal for the black Ternus and red Ternus,  $F(1,16) = 1.355$ ,  $p = 0.262$ . There was no significant difference between PSEs across pre-test and post-test,  $F(1,16) = 0.841$ ,

$p = 0.373$ . No interaction between the two factors was found either,  $F(1,16) = 1.272$ ,  $p = 0.276$ . These results are shown in Figure 4.

The JNDs were  $34.4 \pm 3.9$  (SE) and  $36.6 \pm 4.2$  for the pre-test of the black Ternus type and the pre-test of the red Ternus type. JNDs were  $38.9 \pm 5.2$  and  $31.0 \pm 3.8$  for the post-test of the black Ternus type and the post-test of the red Ternus type. A repeated measures ANOVA with independent factors [Ternus type: Color (black or red) vs. Time (pre-test or post-test)] showed that the JNDs were statistically equal for the black Ternus and the red Ternus,  $F(1,16) = 1.783$ ,  $p = 0.200$ . There was no significant difference between JNDs across the pre-test and post-test,  $F(1,16) = 0.105$ ,  $p = 0.750$ . Importantly, however, the interaction between the factor of Ternus type and Time was significant,  $F(1,16) = 10.034$ ,  $p < 0.01$ . Simple main effects analysis showed that after training with the VAAV temporal structure, the JNDs for the black visual Ternus motion were increased,  $F(1,16) = 5.32$ ,  $p < 0.05$ . In contrast, after training with the AVVA temporal structure, the JNDs for the red visual Ternus motion were decreased,  $F(1,16) = 4.67$ ,  $p < 0.05$ .

Used as a control for Experiments 1a,b showed PSEs of  $121.9 \pm 8.1$  (SE) and  $126.4 \pm 6.2$  for the pre-test of the black Ternus type and the pre-test of the red Ternus type. PSEs were  $120.6 \pm 7.8$  and  $121.8 \pm 9.3$  for the post-test of the black Ternus type and the post-test of the red Ternus type. A repeated measures ANOVA with independent factors [Ternus type: Color (black or red) vs. Time (pre-test or post-test)] showed that the PSEs were statistically equal for the black Ternus and the red Ternus,  $F(1,10) = 1.017$ ,  $p = 0.337$ . There was no significant difference between PSEs across the pre-test and post-test,  $F(1,10) = 0.538$ ,





**FIGURE 4 | PSEs and JNDs for Experiment 1 and Experiment 2.** Experiment 1: “Black” (Experimental Condition for the black Ternus frames), “Black-Control” (Control Condition for the black Ternus frames), “Red” (Experimental Condition for the red frames of the Ternus display), “Red-Control” (Control Condition for the red Ternus frames). Experiment 2: “RB-BR”: Consecutively presented two visual frames of red-black and black-red disks. “RR-BB”: Consecutively presented two visual frames of red-red and black-black disks. The error bars represent standard errors. An asterisk (\*) indicates a significant effect ( $p < 0.05$ ).

$p = 0.480$ . The interaction between the factors of Ternus type and Time was also not significant,  $F(1,10) = 0.288$ ,  $p = 0.603$ .

For Experiment 1b, the JNDs were  $42.5 \pm 6.4$  (standard error) and  $41.3 \pm 5.4$  for the pre-test of the black Ternus type and pre-test of the red Ternus type. The JNDs were  $32.4 \pm 6.8$  and  $31.5 \pm 4.3$  for the post-test of black Ternus type and post-test of the red Ternus type. A repeated measures ANOVA with independent factors (Ternus type: Color (black or red) vs. Time (pre-test or post-test)) showed that the JNDs were statistically equal for the black Ternus and the red Ternus,  $F(1,10) = 0.090$ ,  $p = 0.770$ . However, the JNDs were significantly reduced across the pre-test and post-test,  $F(1,10) = 7.554$ ,  $p < 0.05$ . The interaction between the factor of Ternus type and Time was not significant,  $F(1,10) = 0.006$ ,  $p = 0.938$ .

Overall, Experiments 1a and 1b showed that without auditory-visual temporal interval training, the post-test of Ternus motion increased the sensitivity of categorizing between element motion vs. group motion, and that this improvement was probably due to increased familiarity with the task in the post-test. Importantly though, the different statistical training on the VAAV and AVVA temporal structures led to opposite aftereffects (changes in JNDs) on the perception of Ternus apparent motion: the VAAV condition led to decreased sensitivities for Ternus motion perception while the AVVA condition led to the opposite-sharpened sensitivities for Ternus motion perception.

## Training Performance

The mean accuracy of discrimination between auditory intervals and visual intervals for the VAAV temporal structures was 89.6% (2.7%). The mean accuracy for the AVVA temporal structures was 86.6% (3.6%). Thus, the accuracy rate in the VAAV condition was higher than the one in the AVVA condition,  $t(16) = 3.071$ ,  $p < 0.01$ . However, both accuracy rates were above the chance level of 50%,  $ps < 0.001$ . The results indicate that the training task was successful.

## EXPERIMENT 2

In Experiment 2, we broke up the paired visual and auditory events and presented only a single auditory-visual pair in each trial. The training task was audio-visual TOJ. We examined whether and how the intersensory binding of temporal orders of auditory and visual events bias responses in the post-test of Ternus motion. Although a single sound has not been shown to be potent enough to influence visual apparent motion (Bruns and Getzmann, 2008; Shi et al., 2010), we hypothesized that through intersensory binding, perceptual grouping of dominant temporal structures could still occur, and be used to influence the subsequent perception of visual Ternus motion.

## Participants

Fifty-eight students (35 females) from Peking University took part in Experiment 2. The mean age of the sample was 21.6 years old. The participants were separated into four groups and each group performed only one of the following experimental tasks. For Experiment 2a, there were 15 participants (12 females) with a mean age of 20.4 years old. This group received the “RB–BR” Ternus configuration with interim training on TOJ. Experiment 2b constituted the control group for Experiment 2a. In Experiment 2b, there were 11 participants (six females) with a mean age of 22.4 years old. These participants received the “RB–BR” Ternus configuration without training. In Experiment 2c there were 17 participants (nine females) whose mean age was 21.2 years old. These participants received the “RR–BB” Ternus configuration with interim training on TOJ. Experiment 2d constituted the control group for Experiment 2c. In Experiment 2d there were 15 participants (eight females) whose mean age was 21.3 years old. These participants received the “RR–BB” Ternus configuration without training.

All participants had normal or corrected-to-normal vision and normal hearing. All were naïve as to the purpose of the experiment. The study was approved by the Ethics Committee of the Department of Psychology at Peking University and informed consent was obtained before the experiment for all participants.

## Stimuli and Apparatus

Stimuli and apparatus were the same as those in Experiment 1, except that the color of the disks was different. Details will be described in the following sections.

## Design and Procedure

Experiment 2 was a 2 (type of Ternus configuration: element motion vs. group motion)  $\times$  2 (test group: experiment vs. control) factorial between-participants design.

In Experiment 2, we used a variant of the Ternus paradigm. The direction of the Ternus apparent motion was always from left to right in Experiment 2. We composed the typical dominant element motion and the typical group motion percepts based on the color feature bindings of the Ternus component disks (Kramer and Yantis, 1997; Petersik and Rice, 2008). In the typical element motion condition, in each frame and from left to right, the first Ternus frame contained red and black disks, while the second Ternus frame was composed of black and red disks (referred to as “RB” and “BR” frames). This configuration gave rise to the dominant percept of element motion. In the typical group motion configuration, the first frame was composed of two red disks while the second frame was composed of two black disks (referred to as “RR” and “BB” frames).

## Pre-test and Post-test

In Experiment 2, the settings of the demo and practice, as well as the procedure of the pre-test and post-test were the same as those in Experiment 1, except that the apparent motion direction of the Ternus display was always from left to right. For Experiments 2a and 2b, participants were required to discriminate between element motion and group motion based on the RB–BR Ternus configuration. For Experiments 2c and 2d, participants were

asked to discriminate between element motion and group motion based on the RR–BB Ternus display. In both Ternus settings, the ISI between the two visual frames was from 50 to 230 ms (with 30 ms as a step size).

## Training

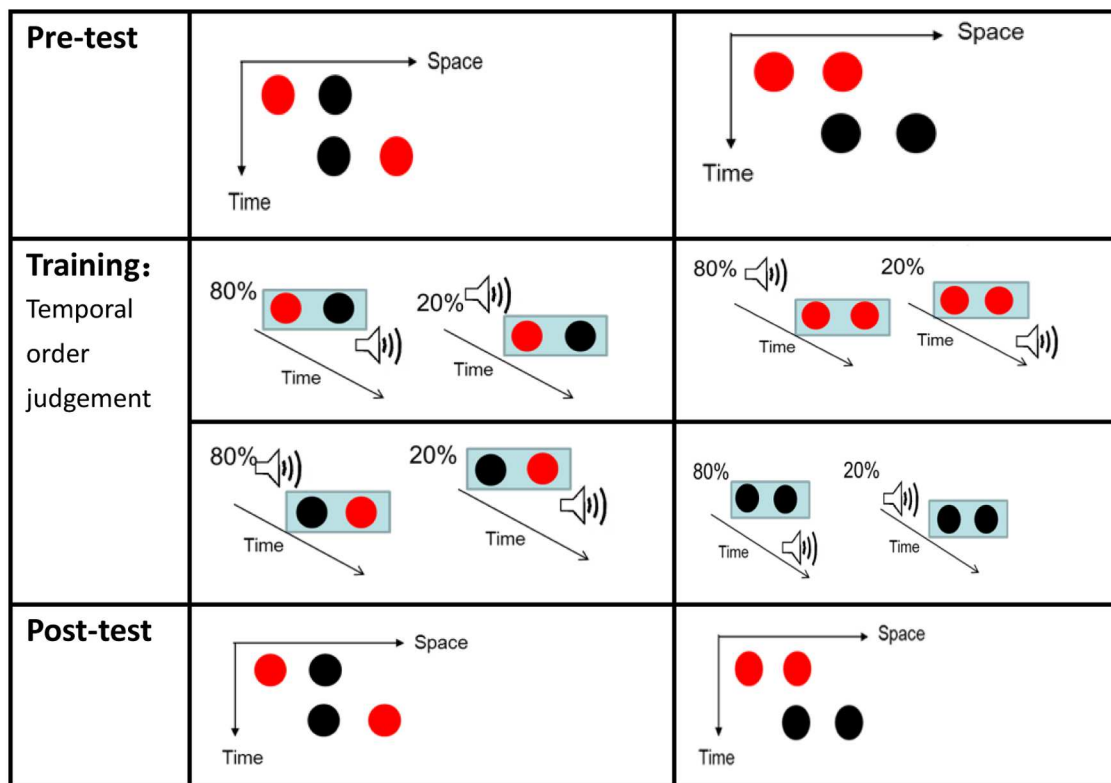
In the training phase, participants performed a TOJ task. After the appearance of a fixation cross (for 300 ms) and a blank display (for 300–500 ms), a visual frame as well as a sound beep appeared with a random stimulus onset asynchrony (SOA) of 50–150 ms. When the presentation of audiovisual stimuli was over, participants were prompted to indicate which stimulus came first. Half of the group was required to press the left mouse key if they perceived the beep first, and the right key if the visual frame was first. The other half of the group reversed the mapping between the response and the stimuli. The inter-trial interval (ITI) was 800–1000 ms. Upon each incorrect response, the word “Wrong” appeared in red on the center of the screen, to give accuracy feedback to the participants. There were 360 trials in total and the training session was separated into six blocks. Participants were asked to take a rest between blocks.

The temporal disparities between auditory–visual pairs and the statistical distribution of those disparities were as follows. In the RB–BR configuration (i.e., the typical element motion percept), 80% of the RB frames preceded the sound beep and 80% of the BR frames trailed the beep. In contrast, 20% of RB frames trailed the sound beep and 20% of the BR frames preceded the sound beep. Henceforth the above distribution would lead to a subjectively dominant perceptual grouping of VAAV. In the RR–BB configuration (typical group motion percept), 80% of the RR frames trailed the sound beep and 80% of the BB frames preceded the beep. In contrast, 20% of RR frames preceded the sound beep and 20% of the BB frames trailed the sound beep. Henceforth the above distribution would lead to a subjectively dominant perceptual grouping of AVVA. This pattern is shown in Figure 5.

## Result

We first conducted a one-way ANOVA of the PSEs and JNDs in the pre-test across the four experiments. The results were as follows: for the PSEs,  $F(3,57) = 1.284$ ,  $p = 0.289$ . For the JNDs,  $F(3,57) = 1.071$ ,  $p = 0.369$ . These results indicate that the baselines for the performance on the Ternus motion tasks between the experimental group and control group were comparable. Therefore, we carried out an independent analysis for each sub-experiment.

In Experiment 2a (the element motion configuration), the PSEs were  $145.0 \pm 5.1$  and  $134.9 \pm 5.1$  for the pre-test and post-test,  $t(14) = 2.574$ ,  $p < 0.05$ . The JNDs were  $31.0 \pm 3.0$  and  $21.6 \pm 2.2$  for the pre-test and post-test,  $t(14) = 3.112$ ,  $p < 0.01$ . In Experiment 2b (the element motion control condition), the PSEs were  $135.6 \pm 9.4$  and  $131.6 \pm 4.7$  for the pre-test and post-test,  $t(10) = 0.507$ ,  $p = 0.623$ . The JNDs were  $38.0 \pm 4.2$  and  $30.5 \pm 3.1$  for the pre-test and post-test,  $t(10) = 2.692$ ,  $p < 0.05$ . Therefore, the training and intersensory binding VAAV temporal structure led to a decreased PSE and more dominant perception of group motion. This is shown in Figure 6.



**FIGURE 5 | Ternus displays (for the pre-test and post-test) and illustrations of the stimuli configurations for Experiment 2.** Two kinds of Ternus displays were used in Experiment 2. The middle of the left side of the figure depicts the training session with the Red-Black Black-Red (RB-BR) configuration. In 80% of the trials here, the Red-Black frame preceded the sound beep and the Black-Red frame trailed the beep. In contrast, in 20% of these trials, the Red-Black frame trailed the sound beep and the Black-Red frame preceded the sound beep. The middle of the right side of the figure depicts the Red-Red Black-Black (RR-BB) frame configuration. In 80% of these trials, the Red-Red frame trailed the sound beep and the Black-Black frame preceded the sound beep. In contrast, in 20% of these trials, the Red-Red frame preceded the sound beep and the Black-Black frame trailed the sound beep.

In Experiment 2c (the group motion configuration), the PSEs were  $135.6 \pm 5.5$  and  $146.8 \pm 4.1$  for the pre-test and post-test,  $t(16) = -2.578$ ,  $p < 0.05$ . The JNDs were  $31.8 \pm 2.6$  and  $24.9 \pm 1.7$  for the pre-test and post-test,  $t(16) = 4.834$ ,  $p < 0.001$ . In Experiment 2d (the group motion control condition), the PSEs were  $149.5 \pm 5.5$  and  $151.9 \pm 4.1$  for the pre-test and post-test,  $t(14) = -0.792$ ,  $p = 0.442$ . The JNDs were  $31.2 \pm 2.3$  and  $28.9 \pm 3.4$  for the pre-test and post-test,  $t(14) = 0.905$ ,  $p = 0.381$ . Therefore, the training and intersensory binding AVVA temporal structure led to an increased PSE and more dominant perception of element motion.

For the training sessions in both experiments, the accuracy of reporting temporal order was better than the chance level. The correct rate was  $95.2 \pm 1.2\%$  for the AV TOJ training in Experiment 2a and  $95.6 \pm 0.8\%$  in Experiment 2c. Therefore, the performance of TOJ was satisfactory.

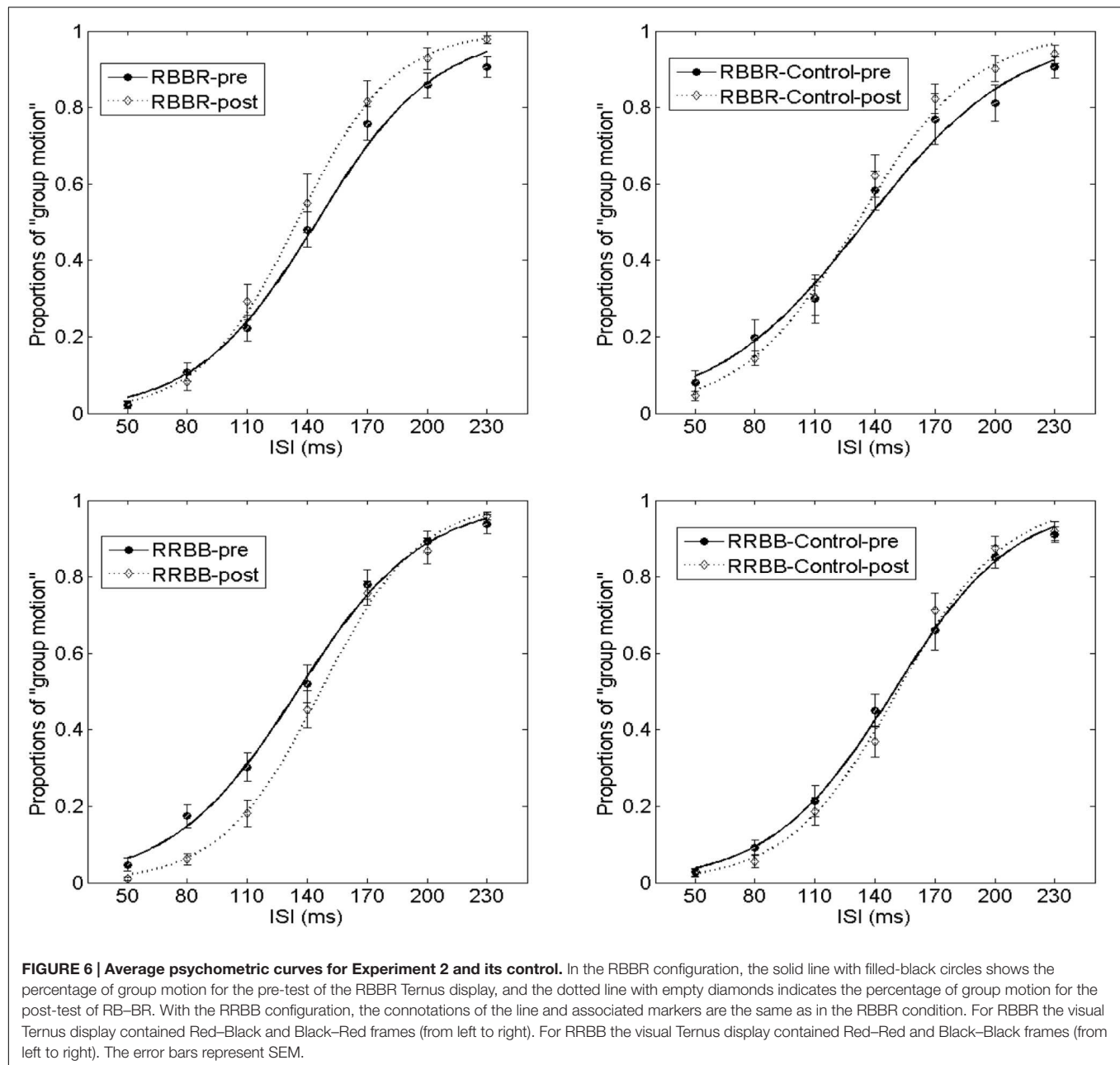
## DISCUSSION

The present study used variants of visual Ternus displays to examine whether fast crossmodal statistical binding of temporal information and sensory properties would recalibrate and hence

bias the perception of visual Ternus motion (element motion vs. group motion). This was indexed by the changes in the PSEs and JNDs of the post-test Ternus apparent motion, compared with the pre-test ones. To achieve this, we manipulated the pairing of audiovisual events in the training session, so that one temporal structure would be statistically dominant over the other. The training tasks were either temporal interval comparisons (Experiment 1) or TOJs (Experiment 2). Therefore, the aftereffects of the training were mainly due to crossmodal statistical binding of temporal information and stimuli properties in the training session.

## Intersensory Binding of Temporal Structure and Audiovisual Properties

In Experiment 1, we created the VAAV temporal structure by putting two black Ternus frames outside of one pair of auditory beeps 80% of the time, so that in this case, the visual interval was longer than the auditory interval. We composed the AVVA temporal structure by putting two red Ternus frames temporally inside the paired auditory beeps 80% of the time, so that the visual interval was shorter than the auditory intervals. According to the temporal precision hypothesis, the auditory interval should



calibrate the visual interval (Burr et al., 2009; Shi et al., 2010). With the VAAV condition, the inserted paired beeps would pull the two black visual frames closer in time, leading to increased JNDs (lower sensitivities) for judging Ternus motion in the post-test. Using the same reasoning, in the AVVA configuration, the flanked (outside) two beeps would pull the two red Ternus frames further away in time, leading to a subjectively extended interval between the two visual frames, and hence higher sensitivities for discriminating Ternus motion. Indeed, from the obtained JNDs, we observed increased sensitivities (smaller JNDs) for Ternus motion in the AVVA condition and decreased sensitivities (larger JNDs) for Ternus motion in the VAAV condition. Compared with the null changes of the PSEs and JNDs in the control tests (only

the pre-test and post-test of Ternus apparent motion), the results from the training protocol suggest that the intersensory binding of the temporal structure as well as audiovisual properties had generalized to affect perception of the implicit 'interval' between the two visual frames in the Ternus display. This led to the perceptual decision of element motion versus group motion.

### Intersensory Binding across Space, Time, and Experimental Trials

One might argue that in Experiment 1 the symmetric alignments of paired audiovisual events on both ends could potentially induce a response bias. This response bias might cause



participants to make TOJs solely on the time asynchrony cues of either the first audiovisual pair or the second audiovisual pair, due to ignorance of the temporal intervals between auditory and visual events. To rule out this possibility, and to provide a comparison for the manipulation in Experiment 1, we directly examined the intersensory binding of visual property and auditory signal when the audiovisual pair was always presented alone. This ensured it would be task-demanding for participants to form the potential interval comparisons between visual pair and auditory pair through the combinations of the onsets and offsets of the two consecutively presented single audiovisual pairs. This was in contrast to Experiment 1, in which two audio visual pairs were always concurrently present, so that intersensory binding between visual property (color) and the auditory beeps (and hence the temporal intervals) occurred spontaneously and was less task-demanding.

Note that in Shi et al. (2010), the authors showed that the transitional threshold for visual apparent motion can be shifted only when the two sounds are presented in temporal proximity to the two visual frames. Such effects were not evident with single-sound configurations. This result suggests that two sounds are required to induce the temporal ventriloquism effect. In other words, the sounds influence the perceived time interval rather than the onset or offset time of visual events. In Experiment 2 of the present study, we explored the main factor that influences visual time perception: the onset and offset time, or, the time interval of the auditory stimuli. If the effect occurs because of the onset and offset time of the auditory stimuli, learning the time information associated with a single pair of asynchronous audiovisual stimuli would be enough to induce the temporal ventriloquism aftereffect. We would further expect to observe more perceptions of element motion for the post-test of the RB-BR Ternus configuration, due to the fact that the two inside auditory beeps would draw the RB and BR visual frame closer. By the same token, the post-test of the RR-BB Ternus display should lead to more reports of group motion. However, this was not the case. In the absence of the events just described, we would expect that the interval information (with both onsets and offsets) is responsible for the temporal recalibration effect. Indeed, we found that in Experiment 2, in which intersensory binding occurred. Observers had formed the dominant temporal structures of V(RB)-AA-V(BR) and A-V(RR)V(BB)-A through the inherent intersensory binding between visual frames of different colors, and their spatial locations, as well as the temporal relations with respect to the corresponding sound beeps. They exploited the different audiovisual interval information to calibrate the probe-test of Ternus apparent motion.

## Mechanisms of ‘Lag Adaptation’ vs. ‘Bayesian Calibration’

As observed in Experiment 2, a “positive” interval adaptation aftereffect contributed to the differing results obtained: (1) with the VAAV temporal structure, the interval between paired visual frames was subjectively extended and led to more reports of group motion (with reduced PSEs) in the post-test; (2) with the

AVVA structure, the interval between the paired visual frames was contracted subjectively and led to more reports of element motion (with increased PSEs) This replicated the results of Zhang et al. (2012). Both temporal structures have led to increased sensitivities of discriminating Ternus motion in the post-test. This positive aftereffect was analogous to the “lag adaptation” revealed in Fujisaki et al. (2004) and other studies, in which they reported that after exposure to a fixed audiovisual time lag for several minutes, human participants showed shifts in their subjective simultaneity responses toward that particular lag (Heron et al., 2007, 2012; Hanson et al., 2008; Harrar and Harris, 2008; Roseboom and Arnold, 2011; Machulla et al., 2012).

In contrast to the findings of Experiment 2, the results from Experiment 1 showed a pattern of the Bayesian calibration (negative) aftereffects (Miyazaki et al., 2006; Yamamoto et al., 2012). In Bayesian negative calibration, when the temporal asynchronies between crossmodal events were sampled from a prior probability distribution (usually a Gaussian distribution), exposure to the above asynchronies led to opposite perceptual changes, and conformed to predictions derived from Bayesian integration theory. In our case, the exposure of subjectively extended “Black-Black” Ternus frame intervals led rather to increased JNDs, and the opposite occurred for “Red-Red” Ternus frames in the post-tests. We inferred that the Bayesian calibration was always at work in both Experiment 1 and Experiment 2. Lag adaptation is advantageous for adjusting variable sound delays that exist in the real world and has been shown to operate in (single) audiovisual pair integration. Thus we speculate that in Experiment 2, the lag adaptation mechanism counteracted the shift of PSEs due to Bayesian calibration, particularly in the judgment of audiovisual intervals, and played an upper hand in determining the temporal aftereffects. This mode of competition between “Bayesian calibration” and “lag adaptation” has also been shown in Yamamoto et al. (2012).

## Implicit Statistical Binding of Temporal Information and Stimuli Properties

Statistical learning, as a theoretical construct, was offered as a general mechanism for learning and processing any type of sensory input that unfolds across time and space (Frost et al., 2014). The present study applied different levels of statistical learning and indicated domain-general ability as well as stimulus-specific constraints in the learning. The learning here refers to updating the internal temporal (interval) representation of the given crossmodal input and encoding potential temporal relations between them. In this way, improvement occurs in the processing of that input and transfers to the post-test of implicit time perception. Here, it manifests as implicit perception of the time interval in the Ternus display. This approach to learning has been investigated explicitly in Roseboom and Arnold (2011). They showed that humans can form multiple concurrent estimates of appropriate timing for audiovisual synchrony, and that audiovisual temporal recalibration can be specific for particular audiovisual pairings. Specifically, participants in Roseboom and Arnold (2011) were shown alternating movies of male and female actors containing positive and negative

temporal asynchronies between the auditory and visual streams. The authors found that audiovisual synchrony estimates for each actor were shifted toward the preceding audiovisual timing relationship for that actor and that such temporal recalibration occurred in positive and negative directions concurrently (Roseboom and Arnold, 2011). Here we found that in addition to forming fixed temporal asynchrony between audiovisual pairings, this intersensory binding could be exploited implicitly by exposure to the more frequent pairing of audiovisual events, but with random temporal asynchronies and correspondence of sensory properties. The intersensory binding in our case may happen more automatically since in the “learning” session, the Ternus frames themselves received less attention than the temporal relations between visual Ternus and auditory beeps. By using an implicit time perception paradigm (Ternus display), the present study has shown a new type of temporal recalibration as a result of comprehensive intersensory binding across time, space, and other sensory properties (such as visual color).

Despite the potential contributions just described, we acknowledge several limitations in the current study. There is an underlying assumption that sound can influence the time perception of visual events because of greater temporal precision with auditory events. However, we did not explore the potential influence of visual events upon auditory time perception in general, nor when the auditory information is blurred. A direction for future research is to explore whether, with appropriate auditory probes, visual temporal information will dominate auditory information in calibrating the auditory timing

task. And although we established the competitive advantages of the Bayesian calibration aftereffects (Experiment 1) and Lag-like adaptation aftereffects (Experiment 2), we did not explore how the specifics of their dominance affect the interpretation of the post-tests and underlying neural mechanisms. This too, is a worthy endeavor for future research.

In summary, by using the typical Ternus effect paradigm, the present study examined crossmodal binding between visual and auditory events across the properties of space, color, and time. The results indicated that depending on the specific binding protocols, statistical binding of temporal information and stimuli properties can concurrently and selectively recalibrate the implicit-time perception of visual intervals. Thus, this binding can influence the perceived states of visual apparent motion.

## AUTHOR CONTRIBUTIONS

LC and YZ designed the experiments. YZ conducted the experiments. LC and YZ analyzed the data and wrote the manuscript.

## ACKNOWLEDGMENTS

This study was supported by grants from the Natural Science Foundation of China (31200760, 81371206) and the National High Technology Research and Development Program of China (863 Program; 2012AA011602).

## REFERENCES

- Brainard, D. H. (1997). The psychophysics toolbox. *Spat. Vis.* 10, 433–436. doi: 10.1163/156856897X00357
- Bruns, P., and Getzmann, S. (2008). Audiovisual influences on the perception of visual apparent motion: exploring the effect of a single sound ☆. *Acta Psychol. (Amst)* 129, 273–283. doi: 10.1016/j.actpsy.2008.08.002
- Burr, D., Banks, M. S., and Morrone, M. C. (2009). Auditory dominance over vision in the perception of interval duration. *Exp. Brain Res.* 198, 49–57. doi: 10.1007/s00221-009-1933-z
- Chen, L., and Vroomen, J. (2013). Intersensory binding across space and time: a tutorial review. *Atten. Percept. Psychophys.* 75, 790–811. doi: 10.3758/s13414-013-0475-4
- Chen, L., and Zhou, X. (2014). Fast transfer of crossmodal time interval training. *Exp. Brain Res.* 232, 1855–1864. doi: 10.1007/s00221-014-3877-1
- Frost, R., Armstrong, B. C., Siegelman, N., and Christiansen, M. H. (2014). Domain generality versus modality specificity: the paradox of statistical learning. *Trends Cogn. Sci.* 19, 117–125. doi: 10.1016/j.tics.2014.12.010
- Fujisaki, W., Shimojo, S., Kashino, M., and Nishida, S. (2004). Recalibration of audiovisual simultaneity. *Nat. Neurosci.* 7, 773–778. doi: 10.1038/nn1268
- Hanson, J. V., Heron, J., and Whitaker, D. (2008). Recalibration of perceived time across sensory modalities. *Exp. Brain Res.* 185, 347–352. doi: 10.1007/s00221-008-1282-3
- Harrar, V., and Harris, L. R. (2008). The effect of exposure to asynchronous audio, visual, and tactile stimulus combinations on the perception of simultaneity. *Exp. Brain Res.* 186, 517–524. doi: 10.1007/s00221-007-1253-0
- Heron, J., Roach, N. W., Hanson, J. V., McGraw, P. V., and Whitaker, D. (2012). Audiovisual time perception is spatially specific. *Exp. Brain Res.* 218, 477–485. doi: 10.1007/s00221-012-3038-3
- Heron, J., Whitaker, D., McGraw, P. V., and Horoshenkov, K. V. (2007). Adaptation minimizes distance-related audiovisual delays. *J. Vis.* 7, 5.1–5.8.
- Ivry, R. B., and Schlerf, J. E. (2008). Dedicated and intrinsic models of time perception. *Trends Cogn. Sci.* 12, 273–280. doi: 10.1016/j.tics.2008.04.002
- Kramer, P., and Yantis, S. (1997). Perceptual grouping in space and time: evidence from the Ternus display. *Percept. Psychophys.* 59, 87–99. doi: 10.3758/BF03206851
- Machulla, T. K., Di Luca, M., Froehlich, E., and Ernst, M. O. (2012). Multisensory simultaneity recalibration: storage of the aftereffect in the absence of counterevidence. *Exp. Brain Res.* 217, 89–97. doi: 10.1007/s00221-011-2976-5
- Miyazaki, M., Yamamoto, S., Uchida, S., and Kitazawa, S. (2006). Bayesian calibration of simultaneity in tactile temporal order judgment. *Nat. Neurosci.* 9, 875–877. doi: 10.1038/nn1712
- Morein-Zamir, S., Soto-Faraco, S., and Kingstone, A. (2003). Auditory capture of vision: examining temporal ventriloquism. *Brain Res. Cogn. Brain Res.* 17, 154–163. doi: 10.1016/S0926-6410(03)00089-2
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat. Vis.* 10, 437–442. doi: 10.1163/156856897X00366
- Petersik, J. T., and Rice, C. M. (2008). Spatial correspondence and relation correspondence: grouping factors that influence perception of the Ternus display. *Perception* 37, 725–739. doi: 10.1068/p5900
- Roseboom, W., and Arnold, D. H. (2011). Twice upon a time: multiple concurrent temporal recalibrations of audiovisual speech. *Psychol. Sci.* 22, 872–877. doi: 10.1177/0956797611413293
- Roseboom, W., Kawabe, T., and Nishida, S. (2013). The cross-modal double flash illusion depends on featural similarity between cross-modal inducers. *Sci. Rep.* 3:3437. doi: 10.1038/srep03437
- Roseboom, W., Nishida, S., and Arnold, D. H. (2009). The sliding window of audio-visual simultaneity. *J. Vis.* 9, 1–8. doi: 10.1167/9.12.4
- Shi, Z., Chen, L., and Muller, H. J. (2010). Auditory temporal modulation of the visual Ternus effect: the influence of time interval. *Exp. Brain Res.* 203, 723–735. doi: 10.1007/s00221-010-2286-3

- Spence, C., and Squire, S. (2003). Multisensory integration: maintaining the perception of synchrony. *Curr. Biol.* 13, R519–R521. doi: 10.1016/S0960-9822(03)00445-7
- Sugano, Y., Keetels, M., and Vroomen, J. (2014). Concurrent sensorimotor temporal recalibration to different lags for the left and right hand. *Front. Psychol.* 5:140. doi: 10.3389/fpsyg.2014.00140
- Treutwein, B., and Strasburger, H. (1999). Fitting the psychometric function. *Percept. Psychophys.* 61, 87–106. doi: 10.3758/BF03211951
- Van der Burg, E., Awh, E., and Olivers, C. N. (2013). The capacity of audiovisual integration is limited to one item. *Psychol. Sci.* 24, 345–351. doi: 10.1177/0956797612452865
- Van der Burg, E., Olivers, C. N., Bronkhorst, A. W., and Theeuwes, J. (2008). Pip and pop: nonspatial auditory signals improve spatial visual search. *J. Exp. Psychol. Hum. Percept. Perform.* 34, 1053–1065. doi: 10.1037/0096-1523.34.5.1053
- Yamamoto, S., Miyazaki, M., Iwano, T., and Kitazawa, S. (2012). Bayesian calibration of simultaneity in audiovisual temporal order judgments. *PLoS ONE* 7:e40379. doi: 10.1371/journal.pone.0040379
- Zhang, H., Chen, L., and Zhou, X. (2012). Adaptation to visual or auditory time intervals modulates the perception of visual apparent motion. *Front. Integr. Neurosci.* 6:100. doi: 10.3389/fnint.2012.00100

**Conflict of Interest Statement:** 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.

Copyright © 2016 Zhang and Chen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# The Application of Timing in Therapy of Children and Adults with Language Disorders

Elzbieta Szelag<sup>1,2\*</sup>, Anna Dacewicz<sup>1</sup>, Aneta Szymaszek<sup>1,2</sup>, Tomasz Wolak<sup>3</sup>, Andrzej Senderski<sup>4</sup>, Izabela Domitrz<sup>5</sup> and Anna Oron<sup>1</sup>

<sup>1</sup> Laboratory of Neuropsychology, Nencki Institute of Experimental Biology, Warsaw, Poland, <sup>2</sup> University of Social Sciences and Humanities, Warsaw, Poland, <sup>3</sup> Institute of Physiology and Pathology of Hearing, Kajetany, Poland, <sup>4</sup> Children's Memorial Health Institute, Warsaw, Poland, <sup>5</sup> Department of Neurology, Warsaw Medical University, Warsaw, Poland

A number of evidence revealed a link between temporal information processing (TIP) and language. Both literature data and results of our studies indicated an overlapping of deficient TIP and disordered language, pointing to the existence of an association between these two functions. On this background the new approach is to apply such knowledge in therapy of patients suffering from language disorders. In two studies we asked the following questions: (1) can the temporal training reduce language deficits in aphasic patients (Study 1) or in children with specific language impairment (SLI, Study 2)? (2) can such training ameliorate also the other cognitive functions? Each of these studies employed *pre-training* assessment, training application, *post-training* and *follow-up* assessment. In Study 1 we tested 28 patients suffering from post-stroke aphasia. They were assigned either to the temporal training (Group A,  $n = 15$ ) in milliseconds range, or to the non-temporal training (Group B,  $n = 13$ ). Following the training we found only in Group A improved TIP, accompanied by a transfer of improvement to language and working memory functions. In Study 2 we tested 32 children aged from 5 to 8 years, affected by SLI who were classified into the temporal training (Group A,  $n = 17$ ) or non-temporal training (Group B,  $n = 15$ ). Group A underwent the multileveled audio-visual computer training *Dr. Neuronowski*<sup>®</sup>, recently developed in our laboratory. Group B performed the computer speech therapy exercises extended by playing computer games. Similarly as in Study 1, in Group A we found significant improvements of TIP, auditory comprehension and working memory. These results indicated benefits of temporal training for amelioration of language and other cognitive functions in both aphasic patients and children with SLI. The novel powerful therapy tools provide evidence for future promising clinical applications.

**Keywords:** temporal information processing, language, cognitive functions, aphasia, specific language disorder

## OPEN ACCESS

### Edited by:

Lihan Chen,  
Peking University, China

### Reviewed by:

Xiangzhi Meng,  
Peking University, China  
Ping Chen,  
Peking University, China

### \*Correspondence:

Elzbieta Szelag  
e.szelag@nencki.gov.pl

### Specialty section:

This article was submitted to  
Perception Science,  
a section of the journal  
Frontiers in Psychology

**Received:** 05 July 2015

**Accepted:** 26 October 2015

**Published:** 12 November 2015

### Citation:

Szelag E, Dacewicz A, Szymaszek A,  
Wolak T, Senderski A, Domitrz I  
and Oron A (2015) The Application  
of Timing in Therapy of Children  
and Adults with Language Disorders.  
Front. Psychol. 6:1714.  
doi: 10.3389/fpsyg.2015.01714

## INTRODUCTION

One of the foundations in modern neuropsychology is the consistent observation that human speech has a dynamic nature and can be analyzed on different temporal levels (Pöppel, 1997). There is strong evidence supporting the thesis that temporal information processing (TIP) on both milli- and multisecond range is a critical factor for speech reception and expression, and provides an



important insight into how our brains process language in norm and pathology (Szelag et al., 2004; Szelag et al., 2011a). In this report we concentrate on millisecond time range.

Experimental data showed that rapid changes in the speech signal such as formant transitions in stop-consonants (like: p, b, t, d, etc.), as well as the voice-onset-time phenomenon reflecting an asynchrony between the burst and the onset of laryngeal pulsing, are rooted in millisecond TIP. Furthermore, strong evidence indicated that various language deficits in children and adults may be associated with deficient TIP (Fink et al., 2006).

Nowadays, the notion about traditionally distinct language disorders, like aphasia, specific language impairment (SLI) or dyslexia, has been reformulated because of a close link between these syndromes. It has been believed that these impairments have similar underpinnings related to deficient TIP. In light of this evidence, the present report is focused on aphasia following cerebral infarction in adults (Study 1) and SLI in children (Study 2).

Aphasia is a consequence of brain damage and defined as an acquired complex language disorder. As only 25% of aphasic patients have a chance for full restoration of disturbed language functions, a need for new therapy methods is huge.

As indicated in pioneer reports by Efron (1963) and Swisher and Hirsh (1972), TIP deficits were evidenced in aphasic patients, independently of the modality tested (auditory or visual). The study by Wittmann et al. (2004) and Fink et al. (2006) reported that patients with deficient comprehension (Wernicke's aphasia) displayed parallel deficits in sequencing abilities, as compared to non-fluent Broca's patients. More recent reports confirmed deficient timing in parallel to language disability in aphasic patients (Oron et al., 2015, for a recent overview see Szelag et al., 2015). These observations were supported by neuroanatomical data indicating an overlapping of structures controlling millisecond timing and language reception (e.g., Wittmann et al., 2004; Lewandowska et al., 2010). Despite these strong evidence, some authors doubt the association '*disordered timing-disordered language*' (Bailey and Snowling, 2002; Heiervang et al., 2002; Share et al., 2002; Ramus et al., 2003). For example, Stefanatos et al. (2007) or Sidiropoulos et al. (2010) indicated either deficient or intact timing, depending on stimulus presentation procedure.

On the other hand, SLI is manifested in disturbances in normal patterns of language acquisition and delayed development of language reception or/and expression which did not result from sensory, emotional, neurological disorders or environmental factors. Although the traditional view assumed that beside language the level of other cognitive functions in SLI remains within the normal range, the more recent studies indicated important deficits in working memory, attention and executive functions (Gathercole and Baddeley, 1990; Cardy et al., 2010; Henry et al., 2012). Although the etiology of SLI remains unclear, researchers agree that deficient TIP reflecting problems in encoding both verbal and non-verbal auditory information may constitute a core deficit in SLI, at least at some cases. These observations may suggest a common neural mechanism that controls both verbal and

non-verbal auditory processing which may be disordered in these children. The present paper is in line with these studies.

Starting from Tallal and Piercy (1973, 1974) who reported that children with SLI are less efficient in discrimination of both speech and non-speech sounds presented in rapid succession, recent studies confirmed difficulties in temporal order judgment for both auditory and visual tasks (Grondin et al., 2007). The other theories implicate some higher-level difficulties associated with language problems (e.g., Rice and Wexler, 1996) or procedural memory deficits (Ullman and Pierpont, 2005). The ambiguity of theories on the core deficits in SLI may be reflected in the great heterogeneity of this disorder.

To sum up, these examples confirmed the relationship '*timing – language*' and indicated the impaired millisecond timing as a candidate for a core symptom underlying the above two distinct language disorders, i.e., post-stroke aphasia in adults and SLI in children.

Given the neurobehavioral similarities between aphasia and SLI, some experimental studies deal with the implication of these findings in therapy programs. The idea behind such implementation is that improvement of disordered TIP through the specific exercises may result in a transfer of improvement from the trained time domain into the untrained language domain. In the existing literature there are a few training programs aiming at reducing deficits in TIP, language, attention, and short-term memory. For example, Scientific Learning developed a training program "Fast ForWord" (FFW) which is broadly spread into the clinical practice. Some studies showed that the effectiveness of computer-based interventions (e.g., FFW) in amelioration of language skills in children with SLI (Stevens et al., 2008), dyslexia (e.g., Gaab et al., 2007), and language-learning impairment (e.g., Tallal et al., 1996). The application of these trainings induces neuroplastic changes in the neural network (e.g., Hayes et al., 2003; Temple et al., 2003; Heim et al., 2013). In contrast, the other studies revealed a similar improvement after the application of FFW or other computer-based interventions in language-disordered children (e.g., Cohen et al., 2005; Gillam et al., 2008; Given et al., 2008). Inconsistency of existing results is still an open question and may be associated with individual differences in disfluency patterns, assessment of language skills, subject age or a group size in particular studies.

In our pilot study (Szelag et al., 2014) we found that in aphasic patients the application of eight sessions of temporal training during 3 weeks yielded improvement of TIP, moreover, a transfer of improvement from the trained time domain into the language domain which remained untrained during the intervention. It was evidenced in language tests assessing auditory comprehension, phoneme discrimination and voicing–unvoicing contrast detection. Importantly, the control non-temporal training did not improve either TIP or comprehension in any applied test. These data seemed promising with respect to future therapy programs addressed patients suffering from aphasia. Therefore, Study 1 presented here aimed at verification of such training effects on a larger patient sample, using longer intervention with modified parameters which may provide more

massive stimulation. Furthermore, the training benefits were assessed with extended diagnostic procedures, focused not only on timing and language, but also on other cognitive functions which are also temporally segmented in the millisecond domain (Pöppel, 1994; Szelag et al., 2011a) and strongly related to the language skill (Oron et al., 2015).

The commonly occurring language disabilities in children and adults inspired us to extend our prototyping training applied in Study 1 into the complex computerized intervention program *Dr. Neuronowski*®. Such new therapy tool applied in Study 2 was focused not only on amelioration of TIP and language, but also other cognitive functions in which TIP is embedded. They are related to language skill and crucial for child mental activity. To test the universality of temporal training benefits we concentrated in Study 2 on SLI in children which constitutes a totally different type of language disorder than aphasia tested in Study 1, but also associated with declined TIP.

## STUDY 1

### Materials and Methods

#### Participants

Twenty eight patients suffering from aphasia after first-ever stroke (16 male, 12 female), aged between 45.8 and 78.9 years ( $x \pm SD = 61.6 \pm 9.1$  years) took part in the study. All subjects were right-handed (Edinburgh Inventory), native Polish speakers and suffered predominantly from receptive language deficits, including disordered auditory comprehension following cerebral infarction ( $n = 27$ ) or cerebral hemorrhage ( $n = 1$ ). The location of brain damage was evidenced by CT or MRI examination (Figure 1). The mean lesion age was  $x \pm SD = 4.21 \pm 3.1$  months. All participants had normal hearing level verified by screening pure-tone audiometry (audiometer AS 208), using frequencies ranging from 250 to 3000 Hz which covered the frequency spectrum of auditory stimuli presented in Study 1. Apart from stroke they had neither neurological nor psychiatric disorders and reported no history of head injuries or severe systemic diseases. The other exclusion criteria were recurrent stroke, global aphasia with poor verbal contact, poor general health, or participation in other rehabilitation programs during our data collection.

Auditory comprehension deficits were evidenced by seven language tests. The comprehension deficits in all subjects were accompanied by disordered TIP which was indicated by increased thresholds for auditory perception of temporal order (see below for detailed description). Moreover, working memory and attentional resources were tested in all subjects.

It was a blinded randomized controlled study. The participants were randomly assigned into two groups according to age, lesion age, level of comprehension and TIP deficits. The experimental group (Group A,  $n = 15$ ) performed a temporal training, whereas, the control group (Group B,  $n = 13$ ) was assigned to a non-temporal training. Synopsis of the *pre-training* performance in Group A and Group B is given in Table 1. Using *U* Mann–Whitney test, all between-group comparisons in the *pre-training* assessment for the above variables were

non-significant with the exception of one aspect of attention which was evidenced by significantly shorter reaction times (RTs) for alertness ( $p < 0.04$ ), corresponding to better performance in Group B than A (Table 1).

#### Neuroanatomical verification of brain damage and between-group balance

To verify the brain lesions structural data (CT or MRI) were conducted. The neuroanatomical data confirmed that the lesioned area in Group A (Figure 1A) and Group B (Figure 1B) comprised almost the same left hemispheric structures, covering the classical areas engaged in both auditory comprehension and TIP (Wittmann et al., 2004; Szelag et al., 2009; Lewandowska et al., 2010). It may be assumed that Group A and Group B were as matched as possible (Table 1, Figure 1).

#### Ethical Approval

The study protocol was approved by the Bioethical Commission at the Medical University of Warsaw (permission no KB/5/2010), as well as by the Ethic Commission at the University of Social Science and Humanities (permission no 2/III/07-08). The study was conducted according to the Helsinki Declaration; the written informed consent from each participant was obtained prior to the testing.

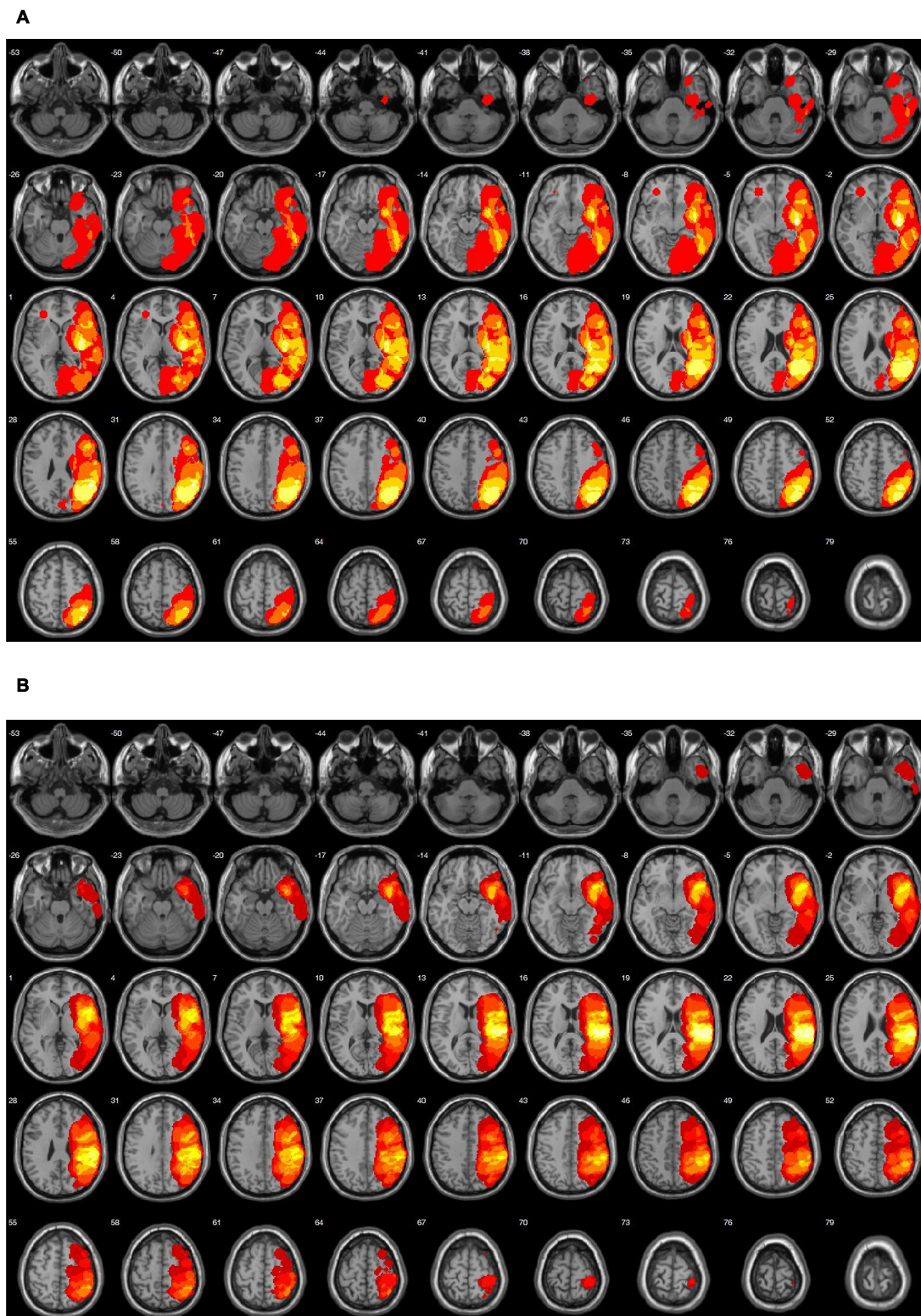
#### Materials and Procedures

The study comprised both assessment and training procedures. The assessment battery included evaluation of TIP, language and other cognitive functions which could influence language skills. The computerized training procedures comprised temporal training (experimental Group A) and non-temporal control training (Group B).

#### Assessment battery

*Assessment of TIP* focused on sequencing abilities in millisecond domain and based on measurement of auditory temporal-order threshold (ATOT; Szymaszek et al., 2009; Szelag and Skolimowska, 2012; Szelag et al., 2014). To summarize briefly, ATOT is defined as the minimum time gap between two auditory stimuli presented in rapid succession that is necessary for a participant to report correctly their order, i.e., the relation *before–after* at 75% correctness. The stimuli were paired 1 ms clicks presented monaurally, i.e., one to each ear with various Inter-Stimulus-Intervals (ISIs). The task was to report the order of these two clicks by pointing to one of two response cards indicating the presentation order: *left–right* or *right–left*. ISI varied adaptively from 1 to 600 ms according to the adaptive maximum-likelihood-based algorithm YAAP (Treutwein, 1997). The ISIs in each trial were set at the current best estimate of the ATOT. This tracking procedure estimates a threshold corresponding to 75% correct order detection based on a logistic psychometric function. The stimulus presentation was terminated when the location of the ATOT was located with a probability of 95% inside a  $\pm 5$  ms interval around the currently estimated threshold.

The stimuli were generated by a 16-bit Sound Blaster Extigy Card and delivered through headphones at a comfortable listening level. Each paired-stimulus was preceded by a warning signal. The proper data collection followed the introductory



**FIGURE 1 | A schematic display of the common lesioned area in Groups A and B in horizontal axis in 20 out of 28 participants. Lesions were shown with yellow/orange color. (A) Group A ( $n = 9$ ): the brightest color indicates the common brain damage in five subjects. It comprises left hemispheric *superior temporal gyrus*, *supramarginal gyrus*, *insula*, *putamen*, *angular gyrus*. (B) Group B ( $n = 11$ ): the brightest color indicates the common brain damage in nine subjects including left hemispheric *superior temporal gyrus* and *supramarginal gyrus*.**



**TABLE 1 | Characteristics of two training groups in pre-training assessment in Study 1.**

Measurement		Group A (n = 15) M (SD)	Group B (n = 13) M (SD)	Significance (U Mann-Whitney test)
<b>Subjects data</b>				
	Age (years)	61.5 (8.3)	61.7 (10.3)	Z = -0.207; ns.
	Lesion age (months)	4.3 (3)	4.2 (3.3)	Z = -0.328; ns.
<b>Assessment</b>				
TIP	ATOT (ms)	179 (53)	166 (53)	Z = -1.196; ns.
Language	Token Test-36 (% of errors)	48.3 (22.6)	50.2 (26.4)	Z = -0.322; ns.
	PDPseudo (% of errors)	21.1 (14.9)	15.4 (9.3)	Z = 1.001; ns.
	PD Words (% of errors)	14 (13.9)	12.5 (11.3)	Z = -0.123; ns.
	PHNoise (% of errors)	19.6 (7.2)	23.4 (10.5)	Z = -0.874; ns.
	PHComp (% of errors)	15.2 (8.9)	17 (9.4)	Z = -0.438; ns.
	IT (% of errors)	55.7 (37.2)	68 (42.3)	Z = -1.300; ns.
	ComprVC (% of errors)	38 (22.1)	43.9 (38.4)	Z = -0.093; ns.
<b>Other cognitive functions</b>				
Working memory	SSP (total score)	4.2 (0.94)	4.5 (1.1)	Z = -0.139; ns.
	SWM (total score)	51.9 (26.4)	44.6 (23.1)	Z = -0.853; ns.
Attention	Alertness (RT in ms)	341 (89)	272 (43)	Z = -2.10; p = 0.04
	Vigilance (RT in ms)	741 (213)	655 (134)	Z = -0.931; ns.

PDPseudo, Phoneme Discrimination for Pseudowords; PDWords, Phoneme Discrimination for Words; PHNoise, Phonemic Hearing in Noise; PHComp, Phonemic Hearing for Compressed Speech; IT, Inflection Trials; ComprVC, Comprehension of Verbal Commands; SSP, Spatial Span; SWM, Spatial Working Memory.

session (described in Szelag et al., 2014). ATOTs were collected in two sessions conducted in consecutive days.

*Outcome measure: the mean ATOT from two sessions (in ms).*

**Assessment of language functions.** *Token Test* (Huber et al., 1983) is sensitive to deficits in auditory comprehension that are one of the main aphasic symptoms. The task was to follow spoken commands of increasing length and complexity. Patients' responses were given either by pointing to, or manipulating with plastic tokens (colored squares and circles of two sizes: big/little), e.g., "Touch the white circle after taking away the yellow square." The whole test consisted of 50 commands given in five sections of increasing complexity.

*Phoneme Discrimination for Pseudowords (PDPseudo):* the task was to decide whether two paired pseudowords were the same or different indicating one of two response cards. The pseudowords differed in consonants, contrasting to place of articulation, fricative, voicing, and nasality, as well as in consonant omission or shifting. The entire test comprised 35 paired pseudowords presented in seven series (75% of pairs were different/25% the same).

*Phoneme Discrimination for Words (PDWords; Nowak-Czerwinska, 1994)* was similar to PDPseudo (see above). The only difference was that paired pseudowords were substituted with paired words. The entire test comprised 64 pairs of words presented in eight series (75% different/25% the same).

*Phonemic Hearing in Noise (PHNoise; modified version of Phonemic Hearing Test, Szelag and Szymaszek, 2006).* The measurement comprised phoneme discrimination in sentences presented via headphones on the background of a cocktail-party noise. Following each sentence presentation, the participant was asked to point to one of two pictures corresponding to that sentence. For example, the sentence "This snake is black"

was presented and the patient pointed to one of two pictures presenting either black snake (in Polish : wąż) or black mustache (in Polish: wąs). The test consisted of 106 trials.

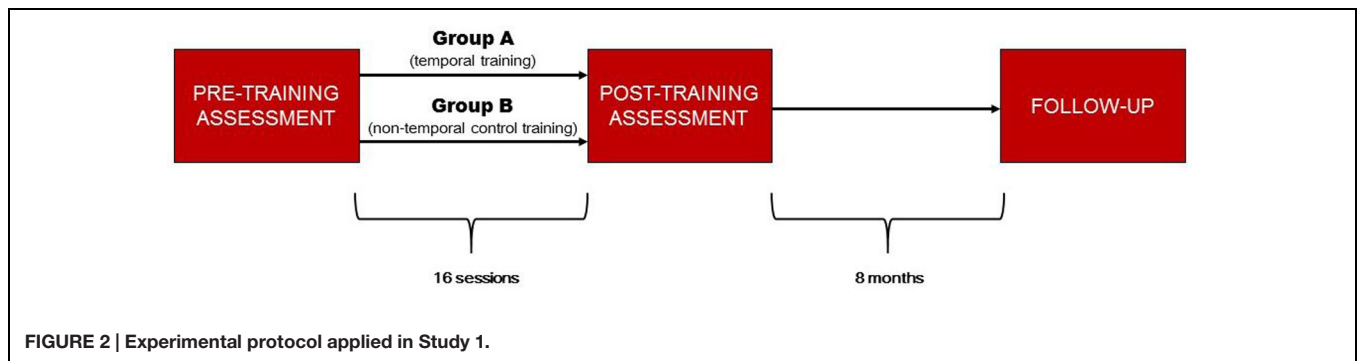
*Phonemic Hearing for Compressed Speech (PHComp; modified version of Phonemic Hearing Test, Szelag and Szymaszek, 2006).* Instead of the background noise, we applied sentences compressed by 20% which made the task more difficult. The experimental protocol was identical as that used for PHNoise.

*Inflection Trials (IT; Lucki, 1995; Szepietowska, 2000)* assessed auditory comprehension of grammar structures using sentences in which one or more grammatical categories with declension (nouns) and suffixes were expressed. In inflection the same word appears in different grammatical forms changing the sentence meaning. The task was to follow spoken commands in six trials, e.g.: "Please, point to the pencil with the key" (In Polish: "Proszę, pokaż ołówek kluczem") and the inflected one: "Please, point to the key with the pencil" (in Polish: "Proszę, pokaż klucz ołówkiem").

*Comprehension of Verbal Commands (ComprVC; Lucki, 1995; Szepietowska, 2000)* assessed the spatial orientation in the body schema, and understanding of the body parts (i.e., forehead, ears, eyes). The task was to perform five commands, e.g.: "Please, point to your nose with the left index finger." *Outcome measures in all language tests: percentage of errors committed in particular tests.*

**Assessment of other cognitive functions.** *Working memory (CANTAB, Cambridge Cognition, 2005).* *Spatial Span (SSP)* measured working memory span in a computerized version of the Corsi Block Test. Various sequences of squares was presented on the screen. The task was to remember the square sequence and to touch the squares in the same order as presented. The number of squares presented within the sequence increased from two to nine. If the subject did not point to the correct sequence





in three consecutive trials, the test was terminated. The longest sequence of squares reflected the span length.

*Outcome measure: the number of memorized items.*

*Spatial Working Memory (SWM)* required retention and manipulation of visuo-spatial information. The task was to find blue ‘tokens’ in presented boxes and use them to fill up an empty column on the right side of the screen.

*Outcome measure: the number of committed errors* (i.e., touching empty boxes and revisiting boxes which already were found to contain a token).

*Attention (Test of Attentional Performance; Zimmermann and Fimm, 1997).* Alertness required simple RTs measurement in response to a visual stimulus (a white cross displayed in the screen center) which in half trials was preceded by an auditory warning signal. The task was to press the response pad after presentation of the cross.

*Vigilance.* A low (440 Hz) and a high (1000 Hz) tone were presented sequentially in a random order during 10-min session. The task was to press a response pad when two identical tones were presented in a row.

*Outcome measure in two attention tests: mean RT achieved in particular test.*

### Training procedures

*Temporal training.* The temporal training used our prototyping procedure developed in previous studies (Szelag et al., 2014). It was rooted in improvement of sequential abilities in the millisecond timing. The main idea of this procedure was complementary to that applied in the assessment of ATOT (see above). Accordingly, the patient was asked to report the order of two clicks presented in rapid sequences with various ISIs. In such a task the shorter ISI corresponded to more difficult task. At the starting point the task difficulty of the training was individually adjusted for each participant on a basis of his/her *pre-training* ATOT (Table 1). The training was provided in 10-trial blocks with fixed ISI in each block. ISIs varied between blocks depending on the score of correct/incorrect responses according to an algorithm based on the patient’s actual ATOT. When at least 90% of correct responses in a block was achieved, the ISI in following block was decreased (increasing task difficulty) according to the following rules: (1) if the actual ATOT was longer than 100 ms, ISI decreased by 5 ms; (2) in case of ATOT between 50 and 100 ms ISI decreased by 2 ms; (3)

for ATOT below 50 ms by 1 ms. In case of correct score within a block below 90%, the ISI increased (decreasing task difficulty) according to the same algorithm as described in (1), (2), and (3).

After each subject’s response a visual feedback on correctness was provided. Additionally, an extra motivation reward system was applied. The participant obtained 1 point after each correct response, 1 point was subtracted in case of incorrect response. After completion of entire 10-trial block the number of collected points was displayed on the screen. Finally, following completion of each block, the participant was rewarded with puzzles which were collected on the monitor during the training session.

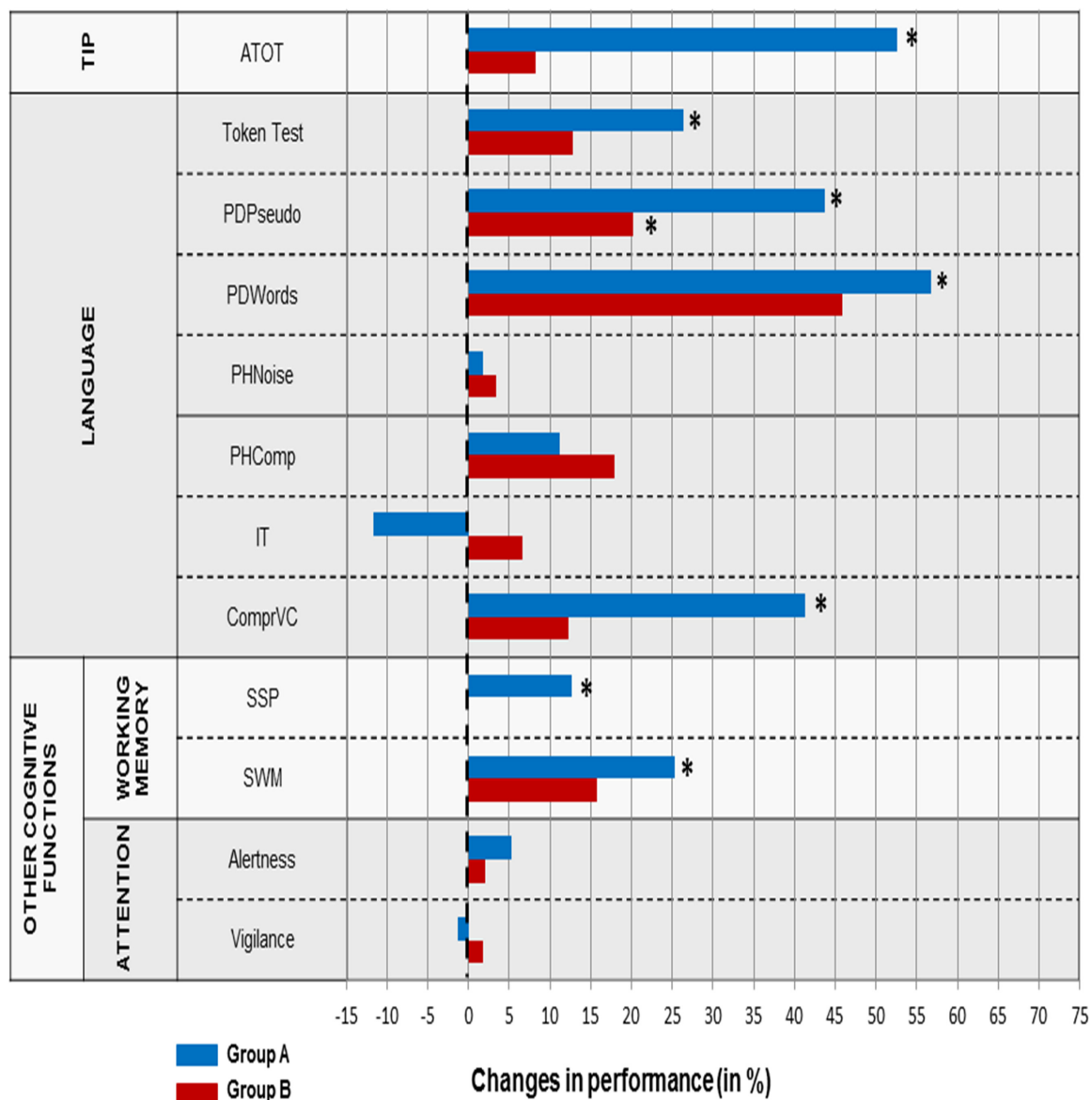
*Non-temporal control training.* The control training was based on loudness discrimination without any aspect of millisecond TIP trained above. Paired 1 s tones separated by a constant ISI of 3 s were presented *via* headphones. One paired tone was always louder than the other tone. The task was to report which tone was louder: the first one or the second one. The adaptivity in the control training based on the loudness difference within paired tones. The loudness of paired tones differs within 0.025–0.00025 of the amplitude range with a constant step of 0.00025. The training was provided in 10-trial blocks. Moreover, four frequencies were used in presented tones: 400, 600, 800, and 1000 Hz, nevertheless within each block only one frequency of presented stimuli was used. In the control training the same motivation system, protocol, as well as comparable mental load was applied as in the temporal training.

### Study protocol

All assessment procedures in each patient were conducted before and after the training (Figure 2). Subsequently, each patient performed individually under supervision of an experimenter 16 sessions of the temporal (Group A) or control (Group B) training for 5 weeks – three sessions per week, each session lasted 45 min. The stability of improvements was verified eight months after the training completion.

### Statistical analyses

To verify the effects of each training type on TIP, language and the other cognitive functions (*pre- vs. post-training* performance), as well as the stability of these effects (*follow-up vs. post-training*) Wilcoxon Signed-Rank test for two dependent samples (within-group comparisons) was performed.



**FIGURE 3 |** Difference in % in the level of *pre- vs. post-training* performance in Study 1 for particular tasks in Group A and B. The dashed vertical line reflects a stable performance (no difference between *pre- and post-training* performance). The positive values (on the right from the dashed vertical line) reflect the improved performance after the training. The minus values (on the left from the dashed line) display the worsened performance. Significant differences ( $p < 0.05$ ) are indicated by asterisks.

## Results

The effect of temporal and non-temporal training was evaluated for particular tasks. The profile of changes in *pre- vs. post-training* performance is given on **Figure 3**.

### Timing, language and other cognitive outcomes following two training types: *pre- vs. post-training* comparisons

#### Temporal Information Processing

**ATOT:** the threshold values *pre-training* ( $\bar{x} = 179$  ms) were significantly higher ( $Z = -3.408$ ,  $p < 0.001$ ) than those

*post-training* ( $\bar{x} = 85$  ms) in Group A. Such difference in Group B was non-significant (166 ms vs. 153 ms in *pre- and post-training*, respectively).

#### Language functions

**Token Test:** the percentage of errors *pre-training* ( $\bar{x} = 48\%$ ) was significantly higher ( $Z = -2.487$ ,  $p < 0.02$ ) than that *post-training* ( $\bar{x} = 35.6\%$ ) in Group A. Such difference Group B was non-significant (50 vs. 43.8% in *pre- and post-training*, respectively).

**PDPpseudo:** the percentage of errors *pre-training* ( $\bar{x} = 21\%$ ) was significantly higher ( $Z = -2.955, p < 0.004$ ) than that *post-training* ( $\bar{x} = 11.9\%$ ) in Group A. The similar relationship ( $Z = -2.355, p < 0.02$ ) was found in Group B (15 vs. 12.6% in *pre-* and *post-training*, respectively).

**PDWords:** the percentage of errors *pre-training* ( $\bar{x} = 14\%$ ) was significantly higher ( $Z = -2.950, p < 0.004$ ) than *post-training* ( $\bar{x} = 6\%$ ) Group A. Such difference in Group B was non-significant (12.5 vs. 6.8% in *pre-* and *post-training*, respectively).

**PHNoise:** the difference in the percentage of errors committed in the *pre-* vs. *post-training* assessment in both groups was non-significant. The mean percentage of committed errors in Group A was 19.6 vs. 19.2% in *pre-* and *post-training*, respectively, whereas in Group B it was 23.4 vs. 22.6%, respectively.

**PHComp:** the difference in the percentage of errors *pre-* vs. *post-training* was non-significant in both groups. The mean percent of errors in Group A was 15.3 vs. 13.5% for *pre-* and *post-training*, respectively and in Group B it was 17 vs. 13.9%, respectively.

**IT:** the difference in the percentage of errors in the *pre-* vs. *post-training* assessment was non-significant in both groups. The mean percentage of errors committed in Group A was 51.7 vs. 57.7% in *pre-* and *post-training*, respectively and in Group B it was 67.9 vs. 63.5%, respectively.

**ComprVC:** the percentage of errors *pre-training* ( $\bar{x} = 38\%$ ) was significantly higher ( $Z = -2.144, p < 0.04$ ) than that *post-training* ( $\bar{x} = 22.3\%$ ) in Group A. Such difference in Group B was non-significant (43.9 vs. 38.5% in *pre-* vs. *post-training*, respectively).

### The other cognitive functions

**Working memory. SSP:** the number of memorized items ( $\bar{x} = 4.2$ ) in *pre-training* assessment was significantly lower ( $Z = -1.999, p < 0.05$ ) than that *post-training* ( $\bar{x} = 4.7$ ) in Group A. Such difference in Group B was non-significant (4.5 vs. 4.5 for *pre-* and *post-training*, respectively).

**SWM:** the number of errors *pre-training* ( $\bar{x} = 51.9\%$ ) was significantly higher ( $Z = -2.358, p < 0.02$ ) than that *post-training* ( $\bar{x} = 38.7\%$ ) in Group A. Such difference in Group B was non-significant (44.6 vs. 37.6% in *pre-* and *post-training*, respectively).

**Attention. Alertness:** the difference in the mean RT achieved in *pre-training* assessment compared to *post-training* assessment was non-significant in both groups. The mean RT in Group A was 341 vs. 323 ms in *pre-* and *post-training*, respectively and in Group B it was 272 vs. 267 ms, respectively.

**Vigilance:** the difference in the mean RT achieved *pre-* vs. *post-training* was non-significant in both groups. The mean RT in Group A was 741 vs. 751 ms for *pre-* and *post-training*, respectively and in Group B it was 655 vs. 643 ms, respectively.

## Stability of Improvements Obtained in Group A

The stability of obtained improvements (*follow-up* vs. *post-training* performance) was assessed only in Group A for tasks in which significant improvement *post-* vs. *pre-training* was evidenced. The *follow-up* assessment in Group A was performed in 10 patients due to the worsened health status in the other ones.

To summarize, the lack of significant differences between the *follow-up* and *post-training* assessment confirmed relatively stable improvements in Group A. It indicated the benefits of temporal training for timing, language and cognitive functions. The obtained outcome measures indicated the continuous progress in subjects' behavior following the temporal training (Table 2).

## Summary of Results

The application of the temporal training in aphasic patients ameliorated significantly TIP which was reflected in lower ATOT values, language skill evidenced in Token Test, PDPpseudo, PDWords, and ComprVC, as well as working memory capacity verified in SSP and SWM. Following the temporal training no significant improvement was reported in PHNoise, PHComp, IT, as well as in Alertness and Vigilance of attention. On the other hand, following the non-temporal training in Group B no significant improvement was observed in any of the applied test except PDPpseudo. All reported improvements were relatively stable for 8 months after the temporal training.

## STUDY 2

## Materials and Methods

### Participants

Thirty two children (10 girls, 22 boys) suffering from SLI (F.80.1 and F.80.2 according to ICD 10; World Health Organization, 1992) aged between 5 and 8 years participated in the study. They were recruited at the Early Intervention Centre and the Children's Memorial Health Institute in Warsaw. All subjects were right-handed (Edinburgh Inventory) and native Polish speakers. The language delay was defined as a reduced performance evidenced by the Test for Assessment of Global Language Skills (TAGLS, Tarkowski, 2001) which constitutes the screening assessment for language development in Polish children. All participants obtained the overall standard score or at least two standard subtests below or equal fourth sten. All children represented the normal level of non-verbal intelligence (IQ at least 85 or higher, measured by the Raven's Colored Progressive Matrices, Szustrowa and Jaworowska, 2003) and normal hearing level (screening pure-tone audiometry, audiometer AS 208) for frequencies ranging from 500 to 4000 Hz which covered the frequency spectrum of auditory stimuli presented in Study 2. The exclusion criteria were neurological, psychiatric, socio-emotional or attentional disorders (as determined by the parental report), as well as the participation in other therapy during our data collection.

It was a blinded randomized controlled study. The recruited children were randomly assigned into two groups (experimental

**TABLE 2 | Summarized results of stability of improvement in Study 1 in Group A follow-up vs. post-training comparisons.**

		Group A				
		<i>n</i>	<i>Post- M (SD)</i>	<i>Follow-up M (SD)</i>	<i>Z</i>	<i>p</i>
TIP	ATOT	6	68 (34)	83 (26)	−1.153	ns.
Language	Token Test	10	32.2 (30)	26.4 (26)	−1.183	ns.
	PDPpseudo	10	12.1 (9.9)	10.9 (8.9)	−1.245	ns.
	PDWords	10	7.5 (16.3)	7.5 (11.7)	−0.282	ns.
	ComprVC	4	22.5 (25)	20 (0)	c	
<b>Other cognitive functions</b>						
Working memory	SSP	7	4.8 (0.6)	5 (0.5)	−1.000	ns.
	SWM	9	39.1 (20.5)	41 (14.8)	−0.178	ns.

As in the follow-up assessment some patients did not perform the full set of tasks, the number of tested participants is given at each task. c, statistical analysis was not performed because of insufficient number of subjects. Values of ATOT are provided by ms; Token Test, PDPpseudo, PDWords and ComprVC by % of errors; SSP and SWM by total score.

**TABLE 3 | Characteristics of two training groups in the pre-training assessment in Study 2.**

Measurement		Group A ( <i>n</i> = 17) <i>M (SD)</i>	Group B ( <i>n</i> = 15) <i>M (SD)</i>	Significance ( <i>U</i> Mann–Whitney test)
<b>Subjects data</b>				
	Age (years)	6.2 (1)	5.9 (0.7)	<i>Z</i> = 0.370; ns.
	IQ	102.9 (30)	113.1 (16.9)	<i>Z</i> = 0.313; ns.
	Language development TAGLS (stems)	2.9 (1.8)	2.7 (1.3)	<i>Z</i> = 0.186; ns.
<b>Assessment</b>				
TIP	ATOT (ms)	196 (6)	211 (76)	<i>Z</i> = −0.672; ns.
Language	Token Test-36 (% of errors)	52.9 (25.2)	59.8 (21.1)	<i>Z</i> = −0.892; ns.
	PDPpseudo (% of errors)	34.3 (13.6)	32.3 (12.4)	<i>Z</i> = −0.546; ns.
	PDWords (% of errors)	18.9 (13)	21 (14.1)	<i>Z</i> = −0.379; ns.
	SSC (% of errors)	34.6 (28.7)	33.3 (24.3)	<i>Z</i> = −0.076; ns.
	COWAT (number of words)	9.6 (5.6)	7.1 (3.5)	<i>Z</i> = −1.404; ns.
<b>Other cognitive functions</b>				
Working memory	SSP (numbers of errors)	9.4 (2.5)	12.3 (6)	<i>Z</i> = −0.022; ns.
	Digit Span (total score)	1.7 (0.8)	2.1 (0.6)	<i>Z</i> = −1.158; ns.
	VWM (total score)	6 (2)	5.3 (1.6)	<i>Z</i> = −0.835; ns.
Attention	Alertness (RT in ms)	438 (155)	481 (161)	<i>Z</i> = −1.322; ns.
	Mazes (total score)	20.4 (6.9)	17.5 (6.0)	<i>Z</i> = −1.331; ns.
Executive functions	TOL <sup>DX</sup> (total move score)	55.6 (19.7)	64.8 (16.9)	<i>Z</i> = −1.559; ns.

TAGLS, Test for Assessment of Global Language Skills; PDPpseudo, Phoneme Discrimination for Pseudowords; PDWords, Phoneme Discrimination for Words; SSC, Syntactic Structure Comprehension; COWAT, Controlled Oral Word Assessment Test; SSP, Spatial Span; VWM, Verbal Working Memory; TOL<sup>DX</sup>, Tower of London Drexel University.

and control) according to age, gender, non-verbal IQ and the level of language development. The experimental group (Group A, *n* = 17) underwent the temporal training and the control group (Group B, *n* = 15) obtained the non-temporal training. Using *U* Mann–Whitney test the two groups did not differ significantly in pre-training assessment for the tested variables (Table 3).

## Ethical Approval

The study protocol was approved by the Bioethical Commission at the Medical University of Warsaw (permission no. KB/162/2010). Written informed consent was obtained from the parents of each child participating in the study, children provided verbal approval.

## Materials and Procedures

Similarly to Study 1, Study 2 comprised both assessment and training procedures. The assessment battery included TIP, language and other cognitive functions e.g., working memory, attentional resources and executive functions. The computerized training procedures comprised temporal training (Group A) and non-temporal control training (Group B).

## Assessment battery

Assessment of TIP based on the same procedure as applied in Study 1, the only difference was that ATOTs were collected from one session only.



**Assessment of language functions.** *Token Test-36* (Kosciesza and Krasowicz, 1995) was a modified version of Token Test for adults used in Study 1 (the whole tests consisted of 30 commands).

**Phoneme Discrimination for Pseudowords (PDPpseudo):** the measurement used the same procedure as applied in Study 1.

**Phoneme Discrimination for Words (PDWords; modified version of Szelag and Szymaszek, 2006):** the measurement used the same procedure as applied in Study 1.

**Syntactic Structures Comprehension (SSC; unpublished materials elaborated in our laboratory):** participants listened to 40 sentences classified into 10 series. Each series contained a set of four sentences of a similar meaning, but differing in either (1) plural vs. singular form or (2) a preposition of place. The task was to indicate on the response card the picture corresponding to one of these four syntactic situations (e.g., "the elephant is standing...in/next to/in front of/ behind...the tent").

**Outcome measures in all above language tests was the percentage of errors committed in particular tests.**

**Controlled Oral Word Assessment Test (COWAT; Lezak, 1995)** measured verbal fluency. The task was to produce as many words as possible from the category of animals during 1 min.

**Outcome measure: the number of produced words.**

**Assessment of other cognitive functions.** *Working memory. Spatial Span (SSP; CANTAB; Cambridge Cognition, 2005)* the same procedure as described in Study 1 was proceeded, nevertheless the outcome measure analyzed here was the number of committed errors.

**Digit Span (Wechsler Intelligence Scale for Children – Revised Version; WISC-R; Matczak et al., 1991):** the task was to listen to the sequence of numbers and recall them in the same order. The number of digits increased from three up to nine. For each correctly reproduced series one point was awarded.

**Outcome measure: total score in the entire test.**

**Verbal Working Memory Test (VWM; unpublished materials elaborated in our laboratory):** the task was to reproduce the order of listened unrelated words presented in series (ranging in words number from two to nine), by pointing to pictures corresponding to presented words. In the first part the words were phonologically similar, whereas, in the second one they were phonologically dissimilar. One point was awarded for each correctly reproduced series.

**Outcome measure: total score in the entire test.**

**Attentional resources.** *Alertness (Zimmermann et al., 2005):* the task was to press a button as fast as possible when the target (picture of a witch presented in a computer screen) appeared in a castle window.

**The outcome measure: the median of RT achieved in the entire test.**

**Executive functions.** *Mazes (WISC-R; Matczak et al., 1991):* the task was to solve 9 maze puzzles of increasing difficulty in a given time limit from 30 to 150 s. For each correctly solved maze, 4

points were awarded. In case of an error one point was subtracted. The error comprised choosing wrong way or passing through the wall.

**Outcome measure: total score in the entire test.**

**Tower of London Drexel University (TOL<sup>DX</sup>; Culbertson and Zillmer, 2005)** consisted of two identical tower structures, one for the subjects and the other for the examiner. Each structure consisted of a board with three pegs and a set of three beads on pegs (red, green, and blue). The task was to replicate 10 configurations of the beads presented by the examiner in as few moves as possible, following two rules: (1) prohibited placement of more beads on a peg than it was accommodated, and (2) to move the beads from pegs one at time.

**Outcome measure: the total move score, reflecting the number of additional moves made while replicating the beads configuration.**

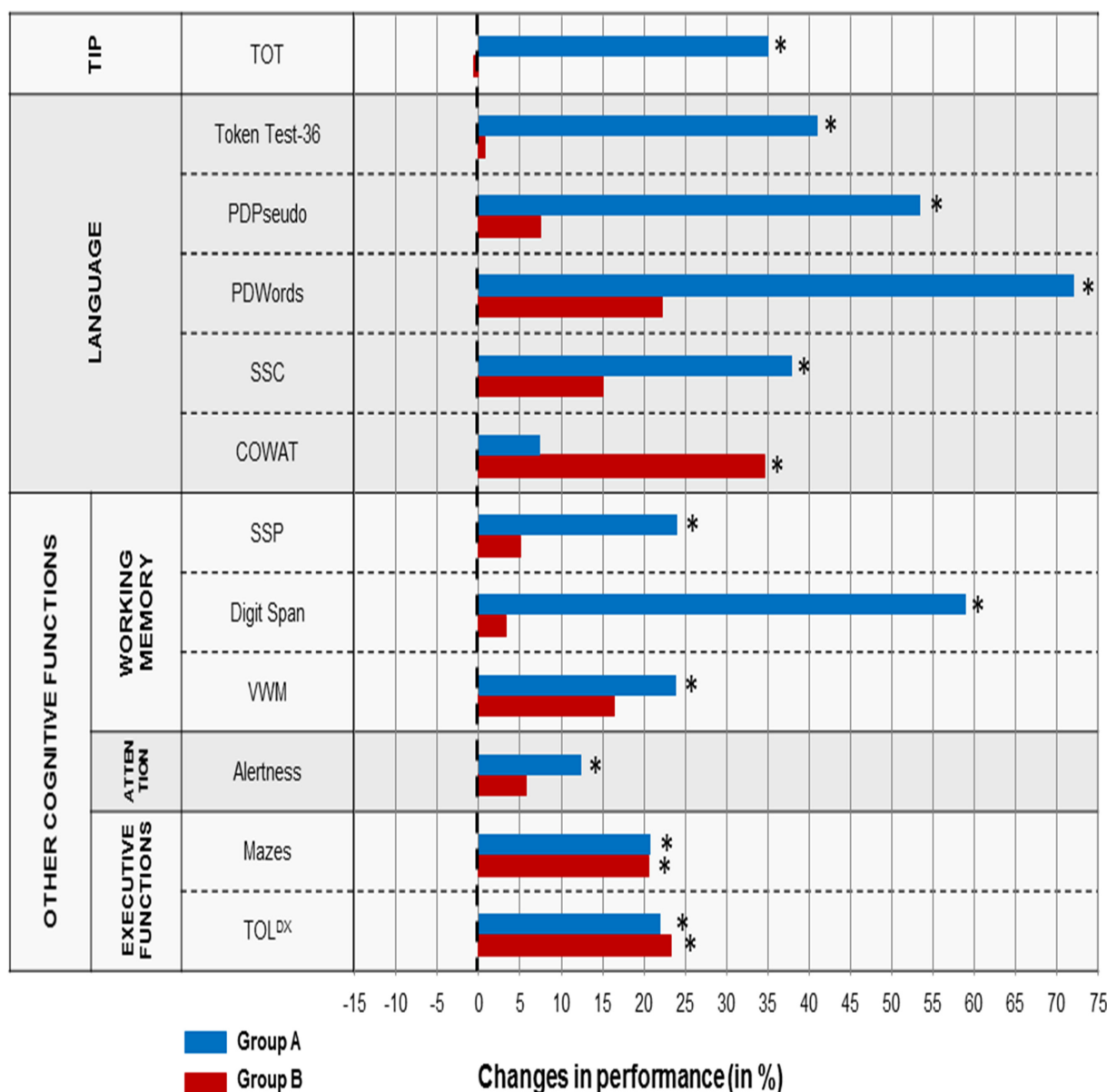
### Training procedures

**Temporal training** procedure used the multimedia intervention program called *Dr. Neuronowski®* ([www.neuronowski.com](http://www.neuronowski.com)) designed in our laboratory. This software consists of nine modules containing 46 games. The majority of these games involved millisecond TIP, sequencing abilities and duration judgment based on results of our previous studies. They were extended by training of other cognitive functions. The task difficulty in particular games changed adaptively on the basis of the actual level of child's performance. The task difficulty comprised: number, length or rate of presented stimuli, various ISIs, the rate of modified speech, application of distractors and time limits for child's responses. The software was designed for tablets, to make it more attractive for children. Particular modules aimed to train the following functions: attention and non-verbal auditory perceptual abilities (*Module 1*), millisecond TIP (*Module 2*, tasks were complementary to both the assessment of ATOT and our prototyping training applied in Study 1), working memory (*Module 3*), executive functions (*Module 4 and 8*), receptive language and phonemic hearing (*Module 5 and 6*), duration judgment of short sounds (*Module 7*), phonemic hearing (*Module 9*) using the Voice-Onset-Time phenomenon (Szelag and Szymaszek, 2014).

**Non-temporal training.** Control training comprised three computer speech-therapy games and 16 computer games available in the Internet (e.g., *Memory or Tetris*), performed on tablets. Speech-therapy games trained phonemic hearing, articulation and vocabulary. The computer games trained attention, working memory and executive functions. Contrary to the temporal training, these tasks did not involve any exercises in rapid auditory processing.

### Study protocol

All assessment procedures were conducted before and after the intervention (**Figure 2**). Each child performed individually, under supervision of an examiner, 24 training sessions of the temporal (Group A) or non-temporal training (Group B) for 6 weeks, four sessions per week, each session lasted 60 min. The stability of improvement was verified



**FIGURE 4 |** Difference in % in the level of *pre- vs. post-training* performance in Study 2 for particular tasks in Group A and B. The dashed vertical line reflects a stable performance (no difference between *pre- and post-training* performance). The positive values (on the right from the dashed vertical line) reflect the improved performance after the training. The negative values (on the left from the dashed line) display the worsened performance. Significant differences ( $p < 0.05$ ) are indicated by asterisks.

in the *follow-up* assessment 6 weeks after the training completion.

## Results

Similarly to Study 1, the effects of each training type on TIP, language and other cognitive functions (*pre- vs. post-training* performance), as well as the stability of these effects (*follow-up vs. post-training*) was assessed using Wilcoxon Signed-Rank test for two dependent samples (within-group comparisons). The profile of changes in

*pre- vs. post-training* performance for each task is given on Figure 4.

## Timing, Language and Other Cognitive Outcomes Following Two Training Types: *Pre- vs. Post-training* Comparisons

### Temporal Information Processing

ATOT values in Group A were significantly higher ( $Z = -3.464$ ;  $p < 0.001$ ) *pre-training* ( $\bar{x} = 196$  ms) than *post-training* ( $\bar{x} = 127$  ms). Such difference in Group B was non-significant

( $\bar{x}$  = 211 ms vs.  $\bar{x}$  = 225 ms for *pre-* and *post-training*, respectively).

### Language functions

**Token Test-36:** the percentage of errors in Group A was significantly higher ( $Z = -3.524$ ;  $p < 0.001$ ) *pre-training* ( $\bar{x}$  = 52.9%) than *post-training* ( $\bar{x}$  = 31.2%). Such difference in Group B was non-significant ( $\bar{x}$  = 59.8% vs.  $\bar{x}$  = 55.6% for *pre-* and *post-training*, respectively).

**PDPseudo:** the percentage of errors in Group A was significantly higher ( $Z = -2.708$ ;  $p < 0.007$ ) *pre-training* ( $\bar{x}$  = 34.3%) than *post-training* ( $\bar{x}$  = 16%). Such difference in Group B was non-significant ( $\bar{x}$  = 32.3% vs.  $\bar{x}$  = 29.9% for *pre-* and *post-training*, respectively).

**PDWords:** the percentage of errors in Group A was significantly higher ( $Z = -3.626$ ;  $p < 0.001$ ) *pre-training* ( $\bar{x}$  = 18.9%) than *post-training* ( $\bar{x}$  = 5.3%). Such difference in Group B was non-significant ( $\bar{x}$  = 21% vs.  $\bar{x}$  = 16.3% for *pre-* and *post-training*, respectively).

**SCC:** the percentage of errors in Group A was significantly higher ( $Z = -2.877$ ;  $p < 0.004$ ) *pre-training* ( $\bar{x}$  = 34.6%) than *post-training* ( $\bar{x}$  = 21.5%). Such difference in Group B was non-significant ( $\bar{x}$  = 33.3% vs.  $\bar{x}$  = 28.3% for *pre-* and *post-training*, respectively).

**COWAT:** the difference between the number of produced words in Group A was non-significant between *pre-training* ( $\bar{x}$  = 9.6) and *post-training* ( $\bar{x}$  = 10.4). In Group B the number of produced words was significantly lower ( $Z = -0.825$ ;  $p < 0.02$ ) *pre-training* ( $\bar{x}$  = 7.1) than *post-training* ( $\bar{x}$  = 9.6).

### The other cognitive functions

**Working memory. SSP:** the number of errors in Group A was significantly higher ( $Z = -2.753$ ;  $p < 0.05$ ) *pre-training* ( $\bar{x}$  = 12.4) than *post-training* ( $\bar{x}$  = 9.4). Such difference in Group B was non-significant ( $\bar{x}$  = 12.2 vs.  $\bar{x}$  = 12.8 for *pre-* and *post-training*, respectively).

**Digit Span:** the total score in Group A was significantly lower ( $Z = -2.640$ ;  $p < 0.008$ ) *pre-training* ( $\bar{x}$  = 1.7) than *post-training* ( $\bar{x}$  = 2.7). Such difference in Group B was non-significant ( $\bar{x}$  = 2.1 vs.  $\bar{x}$  = 2.1 for *pre-* and *post-training*, respectively).

**VWM:** the total score in Group A was significantly lower ( $Z = -2.620$ ;  $p < 0.009$ ) *pre-training* ( $\bar{x}$  = 6) than *post-training* ( $\bar{x}$  = 7.4). Such difference in Group B tended toward significance ( $Z = -1.922$ ;  $p < 0.055$ ;  $\bar{x}$  = 5.3 vs.  $\bar{x}$  = 6.1 in *pre-* and *post-training*, respectively).

### Attention

**Alertness:** median RT in Group A was significantly longer ( $Z = -2.676$ ;  $p < 0.007$ ) *pre-training* ( $\bar{x}$  = 438 ms) than *post-training* ( $\bar{x}$  = 384 ms). Such difference in Group B was

non-significant ( $\bar{x}$  = 481 ms vs.  $\bar{x}$  = 453 ms for *pre-* and *post-training*, respectively).

### Executive functions

**Mazes:** total score in both Group A and Group B was significantly lower ( $Z = -2.648$ ;  $p < 0.008$  and  $Z = -3.182$ ;  $p < 0.001$ , respectively) *pre-training* ( $\bar{x}$  = 20.4 and  $\bar{x}$  = 17.5, respectively) than *post-training* ( $\bar{x}$  = 24.6 and  $\bar{x}$  = 21.1, respectively).

**TOL<sup>DX</sup>:** total move score in both Group A and Group B was significantly higher ( $Z = -2.121$ ;  $p < 0.04$  and  $Z = -2.413$ ;  $p < 0.02$ , respectively) *pre-training* ( $\bar{x}$  = 55.6 and  $\bar{x}$  = 64.8, respectively) than *post-training* ( $\bar{x}$  = 43.5 and  $\bar{x}$  = 49.7, respectively).

### Stability of Changes Obtained in Groups A and B

As following each training type we observed improvements in some tests, the stability of changes was assessed in Groups A and B on the basis of comparisons between *follow-up* vs. *post-training*. Results of these comparisons are given in **Table 4**. Due to reduced subject sample, the *follow-up* assessment was performed in 18 children ( $n = 9$  in each group).

To sum up, the obtained relationships in Study 2 were relatively stable for 6 weeks after training completion. Non-significant differences between *follow-up* vs. *post-training* point to stable training-related changes. They were evidenced in Group A for ATOT, some language tests (Token Test-36, PDPpseudo and PDWords), and other cognitive tests (SSP, Digit Span, VWM, Alertness, Mazes and TOL<sup>DX</sup>). Moreover the percentage of errors in SSC in *follow-up* was significantly lower than *post-training*, indicating continued improvement. Stability of performance was also observed in Group B for COWAT, Mazes and TOL<sup>DX</sup>.

### Summary of Results

The application of the temporal training (Group A) in children with SLI ameliorated significantly TIP which was reflected in lower ATOT values, as well as language skills observed in PDPpseudo, PDWords, SSC and Token Test-36. Moreover, remediated working memory was evidenced in SSP, VWM and Digit Span, as well as attentional resources measured with Alertness. Finally, executive functions were also improved which was evidenced in Mazes and TOL<sup>DX</sup>. Following the temporal training any significant improvement was lacking in COWAT.

## DISCUSSION

Two studies reported here measured the effects of temporal and non-temporal trainings on TIP, language skills and other cognitive functions in two language-disordered groups, i.e., in adult individuals suffering from post-stroke aphasia (Study 1) and in children with SLI (Study 2). Suggestive evidence from these studies indicated a clear dissociation between the beneficial effects of these two major intervention types, i.e., the temporal and non-temporal

**TABLE 4 | Summarized results of stability of changes in Study 2 in Groups A and B for follow-up vs. post-training comparisons.**

		Group A					Group B				
		<i>n</i>	<i>Post-M (SD)</i>	<i>Follow-up M (SD)</i>	<i>Z</i>	<i>p</i>	<i>n</i>	<i>Post-M (SD)</i>	<i>Follow-up M (SD)</i>	<i>Z</i>	<i>p</i>
TIP	ATOT	9	115 (50)	92 (63)	−1.362	ns.	9	221 (63)	193 (89)	−0.140	ns.
Language	Token Test-36	8	38.3 (23.1)	41.3 (26.8)	−0.722	ns.	9	52.6 (17.4)	52.6 (19.3)	0	ns.
	PDPseudo	6	20 (14.7)	20 (14.7)	1	ns.	7	20.8 (14.1)	23.3 (19.4)	−0.813	ns.
	PDWords	9	4.3 (6.4)	3 (5.6)	−1.841	ns.	8	7.5 (8.3)	10 (13.4)	−0.341	ns.
	SSC	8	22.5 (23.3)	15.6 (16.8)	<b>−2.226</b>	<b>0.03</b>	8	18.1 (11)	16.9 (12.8)	−0.424	ns.
	COWAT	7	8.7 (3)	11.1 (4.3)	−1.873	ns.	9	9.1 (5.3)	7.9 (4.3)	−1.199	ns.
<b>Other cognitive functions</b>											
Working	SSP	9	9.1 (2.8)	11.9 (4)	−1.436	ns.	7	12 (6.3)	10.6 (3.2)	−0.170	ns.
memory	Digit Span	5	3 (1)	3.2 (1.1)	−1.000	ns.	3	2 (1)	2.7 (1.2)	c	
	VWM	8	7.1 (1.7)	7.9 (1.7)	−1.511	ns.	6	6.7 (2.2)	7.3 (1.6)	−0.649	ns.
Attention	Alertness	7	370 (53)	400 (85)	−1.352	ns.	8	428 (121)	419 (154)	−0.980	ns.
Executive	Mazes	8	27.5 (6.5)	26.4 (5.2)	−0.853	ns.	7	22.6 (4.2)	22.3 (3.9)	−0.315	ns.
functions	TOL <sup>DX</sup>	8	37.8 (16.1)	33.4 (19.7)	−0.840	ns.	8	47.9 (16.8)	43.9 (15.4)	−0.763	ns.

The significant differences were bolded. As some children in follow-up did not perform the full set of tasks, the number of tested participants is given at each tasks. c, statistical analysis was not performed because of insufficient number of subjects. Values of ATOT are provided by ms; Token Test-36, PDPseudo, PDWords, SSC, SSP by % of errors; COWAT by the number of words; Digit Span, VWM, Mazes by total score and TOL<sup>DX</sup> by total move score.

training. Whereas the former resulted in remediating TIP, language and other cognitive functions, the latter improved only some selected aspects of functions measured in these two studies.

## Increased Efficiency for Rapid Auditory Processing After Temporal Training

Despite a similar initial level of timing performance reflected in non-significant entrance differences in ATOT between Groups A and B in Studies 1 and 2 (see **Tables 1** and **3**), the application of two different training types (temporal or non-temporal) brought about different effects on TIP. The application of temporal training in Groups A (Studies 1 and 2) resulted in significantly lowered thresholds for the order detection *post-* than *pre-training*, corresponding with improved sequencing ability. Such difference proved non-significant in both Groups B (Studies 1 and 2), as reflected in within-group comparisons for *post vs. pre-training* performance (**Figures 3** and **4**). Divergent effects of these two major intervention types may be caused by providing various stimulation based either on rapid auditory processing (temporal training), or on a lack of such stimulation (non-temporal training).

This relationship seems independent of the content of the applied temporal intervention. Using various procedures, durations and protocols in temporal trainings in Studies 1 and 2, we found some important similarities in beneficial effects following their application. In case of aphasic patients (Study 1), we applied our prototyping, rather simple intervention program, focused on sequencing in auditory perception of event order using paired stimuli only. But in children suffering from SLI (Study 2), the more complex stimulation was applied, using the extended exercises and more attractive display. Beside ordering paired stimuli, it comprised more complex sequences of various length and presentation parameters adjusted adaptively.

Moreover, an extra paradigm was implemented focusing on duration judgment in millisecond range. The novel value of these two studies is that two various contents of intervention approaches resulted in improved TIP, moreover, evidenced the transfer of improvement from the time domain into the untrained language and other cognitive domains. Although Study 1 confirms the benefits of the core idea of such therapy, the *Dr. Neuronowski*® program applied in Study 2 seems more drawing a participant into the exercise, hence more attractive for the future users. It seems an optimal method of remediating language and cognitive deficits (detailed discussion below). It may be concluded that the basic content of the intervention is not a matter, but it may be important from a perspective of future users or educators.

Considering the beneficial effects it seems important to indicate a candidate mechanism that may, in concert with others, underlie ameliorated temporal acuity after the intervention. Although the direct evidence cannot be defined yet, one may hypothesize that the millisecond TIP mechanisms, responsible for sequencing abilities may operate on a very fundamental level, regardless the content of the sensory stimulation. To explain one of the possible neural sources of sequencing abilities, one may refer to neuronal spontaneous gamma band oscillations of a periodicity of 40 Hz observed in electrophysiological studies. One period of such oscillatory activity has around 25 ms duration and corresponds in duration with the time range crucial for the perception of temporal order (van Rullen and Koch, 2003). Referring to the idea proposed by Pöppel (2009), the relation ‘before–after’ in incoming rapid events can be detected if two of them occur at least within two successive oscillatory periods. It reflects a situation when the gap between successive events falling in rapid succession is longer than one oscillatory period. The shorter gap creates problems in the proper detection ‘before–after’ relation.



There is a strong experimental support that these neuronal oscillations play an important role in human cognition (Poeppel, 2003; van Rullen and Koch, 2003). To conclude, ameliorated TIP evidenced in lowered ATOT and improved sequencing abilities (Figures 3 and 4) following the temporal training may create a neural basis for remedial gains.

## Divergent Effects of Temporal and Non-temporal Training on Language Skills

The results presented here support the thesis on the close relationship between the millisecond timing and language, which was previously reported in the literature (see Introduction). The deficient millisecond timing, reflected in poorer temporal acuity evidenced in language-disordered population, may overlap with problems in speech perception/expression which is segmented temporally in the time window corresponding in duration with ATOT values.

The results obtained after temporal training in aphasic patients or after *Dr. Neuronowski*® application in SLI children revealed a similar pattern of effects on sequencing abilities and language functions. Independently of the group tested (children, adults), following the temporal training we observed significant improvement of language skills which was not evidenced after the non-temporal intervention (with the exception of COWAT in Study 2). It may support a notion on a fundamental role of temporal acuity in our verbal communication. The novel important finding for clinical practice is the emphasis of the receptive language improvement in two distinct clinical groups following the intervention in millisecond range. It provides an excellent, innovatory and effective tool for the neurorehabilitation of patients suffering from receptive language problems of different etiology.

On the other hand, the novel important finding is that improved temporal acuity may result in enhanced temporal dynamic of information processing, thus, more concert processing with the typical temporal dynamic of the human speech (see Introduction). The optimization within the temporal template may create a neural basis for improved speech perception/expression which is characterized by a typical temporal segmentation. Another possibility would be a contribution of other cognitive functions, like working memory, attention or executive functions which are involved in information processing, including speech processing. These functions might be ameliorated following the intervention providing an enhancement for verbal processing. We discuss these two possibilities below.

The application of the prototypic temporal training in aphasic patients and *Dr. Neuronowski*® intervention in SLI children resulted mainly in amelioration of language processes on phonemic level. These data are consistent with our previous pilot studies in patients with aphasia (Szelag et al., 2014) in which the improvement in receptive language functions was observed even after the simple temporal training. Here, we confirm the significantly better performance in both groups

after various temporal trainings in phoneme discrimination tasks on syllable (PHPseudo), word (PHWords) and sentence level (ComprVC, Token Test). However, no improvement was evidenced in phonemic hearing (PHNoise, PHCompr) and inflection functions (IT).

Phoneme discrimination on syllable and word level measured in Studies 1 and 2 was based on pure auditory processing, without any extra cues. As mentioned in the Introduction, phoneme discrimination uses spectral cues which comprise rapid formant transitions in millisecond time window, similar to that critical for timing on this range. An efficient information processing within this time range is crucial for auditory comprehension on the basic level of phonemes, syllables, and words which constitute the segments of verbal utterances. Thus, improvement of temporal acuity on millisecond level, in other words – speeding up the internal clock or better synchronization of neural oscillations, resulted in more accurate phoneme discrimination. The strong correlation between language and timing was evidenced in our recent study on aphasic patients (Oron et al., 2015). It may be assumed that such temporal mechanism operates on a very basic level, regardless of the kind of material being processed (verbal/non-verbal).

Another important result obtained in both tested groups was the improvement of comprehension of spoken commands of increasing length and complexity (Token Test in Study 1 and Token Test-36 in Study 2). It required not only phoneme discrimination but involved also higher linguistic functions (semantic, syntactic and/or post interpretative processes), moreover, a strong component of visual and auditory processing or working memory load. Following temporal training we observed also better performance in the other tests involving the higher linguistic functions. In patients with aphasia (Study 1) it was evidenced in sentence comprehension (ComprVC) which comprised spatial orientation in the body schema. In children with SLI it was found in syntactic comprehension using SSC. The functions measured in Token Test, Token Test-36, ComprVC and SSC required not only well preserved linguistic processing (phonemic, semantic and syntactic) but also efficient verbal working memory. As the working memory was improved after the temporal trainings in both groups, its contribution to improved performance on these tests would be also possible (see the next section for detailed discussion).

It is interesting to note that in Study 2 in Group B we observed improved verbal fluency and executive functions measured with COWAT. This task required a spontaneous word production, as well as an ability to create, plan and execute activities. The better performance of this task may be explained referring to the strong component in the non-temporal training of expressive language skills (e.g., word presentations and repetitions) which could extended children vocabulary. Moreover, the better performance on COWAT may be associated with overall improvement in Group B in executive functions, measured with Mazes and TOL<sup>DX</sup>. The similar expressive language improvement after the temporal training was observed by Heim et al. (2013).

In aphasic patients after the temporal training non-significant improvement was observed in phonemic hearing using

modified conditions, i.e., a background of noise (PHNoise) and compressed speech (PHComp). Such lack of improvement may be due to the specific test procedure based on pictures displaying the auditory presented sentences. Such extra visual cues might be helpful in phoneme identification. Similarly, in aphasic patients no effect was observed in auditory comprehension of grammar structures, using sentences with internal modifications in which one or more grammatical categories with declension and suffixes were presented (IT). In Polish language these grammar structures constitute the most difficult elements of language which may generate problems even in healthy language users. The constant level of performance with relatively high percentage of errors before and after the training (**Figure 3**) may reflect the task difficulty and the severity of impairment.

To sum up, temporal trainings improve language functions in two distinct clinical groups, suggesting the coexistence of the common neural platform which control information processing. In case of non-temporal training these beneficial effects were generally not found.

## Divergent Effects of Temporal and Non-temporal Training on the Other Cognitive Functions

The positive temporal training effects widespread for the other cognitive functions, i.e., working memory, attention and executive functions. In Study 1 we observed a clear remediation of working memory capacity. After the temporal training patients memorized significantly more items (SSP) and committed significantly fewer errors in SWM. Such improvements may result from improved TIP and higher temporal acuity (see above) which created a modified frame for working memory, as well as other cognitive functions (Szelać et al., 2011b; Bao et al., 2013). Some authors (Ulbrich et al., 2009) emphasized the relations between millisecond timing and working memory. Although in aphasic patients during the temporal training working memory was not trained directly, the changes within a temporal template underlying our mental activity may modify the working memory resources. As a consequence, the remediated working memory may facilitate the performance on other tasks, including some language tasks applied in Study 1 (Token Test, ComprVC) or executive function tasks in Study 2 (Mazes, TOL<sup>DX</sup>) in which the load of working memory is high. A number of existing evidence confirmed the contribution of working memory to cognition, including speech processes (see Baddeley, 2003 for a review). Moreover, to succeed in the training based on temporal ordering, the trained temporal skills had to be accompanied by the efficient working memory load.

Improved working memory was also observed after the application of *Dr. Neuronowski*<sup>®</sup> program in children with SLI (Study 2) which was evidenced in lower errors in SSP, increased total score in VMM and Digit Span (**Figure 4**). As *Dr. Neuronowski*<sup>®</sup> is much more extended tool compared to the intervention applied in aphasic patients, it provided more massive stimulation addressing many cognitive functions directly, including working memory, attention and executive functions.

According to Lezak (1995), attention functions differ from other cognitive functions and they could be treated as mental activity variables which are highly engaged in many other cognitive functions. In children with SLI, decreased RTs in Alertness were observed after the temporal training. In this terms, shorter RT corresponding with better performance, may reflect the increased general processing speed (van Zomeren and Bouwer, 1994). Some authors (Stevens et al., 2008) suggested that improvement in receptive language in children with SLI after administration of FFW may be caused by the enhancement in sustained attention. It should be stressed that the correlation between some aspect of attention and TIP was observed also in our previous studies (Szymaszek et al., 2009; Oron et al., 2015).

Evidence suggest that children with SLI display difficulties in tasks engaging executive functions (e.g., Roello et al., 2015). Alarcón-Rubio et al. (2014) showed that receptive vocabulary skills and self-directed speech usage are associated with executive functions in typically developing children aged from 4 to 7 years. Such relationship may indicate the similar temporal frame provided by the temporal mechanisms on which various cognitive functions are embedded. Such hypothesis finds it support in taxonomy of neuropsychological functions provided by Pöppel (1994) which assumes that TIP provides a logistic basis for many cognitive activities.

In literature data the relationship '*TIP-executive functions*' is a neglected topic. It inspired us to include executive functions into the diagnostic set in Study 2. Interestingly, the improvement of executive functions was evidenced in Mazes and TOL<sup>DX</sup> following the temporal training (Group A) and non-temporal one (Group B). In our opinion, such improvement in Groups A and B may result from different reasons. In Group A it may result from the improvement of the temporal dynamic of the neural network (see above for explanations), direct training of the executive functions (*Module 4* and *8*), or from interaction between these two factors. In contrast, in Group B in Study 2 remediating executive functions might result from playing computer games which were included into some parts of our non-temporal intervention. Its beneficial effects with respect to the mental activity were previously reported by Bavelier et al. (2011).

To conclude, the improved cognition may result from the interrelation between the improved temporal frame and cognitive load contained in the applied intervention in parallel to the TIP exercises. Beneficial effects of temporal and non-temporal training for the mental activity supports the contribution of other cognitive functions to speech therapy.

## Implications for Actual Practice and Future Research

The current report has important implications for clinical practice and future experimental studies. The novel value of two studies presented here is the indication for the first time that two distinct language disorders of various etiologies, i.e., post-stroke aphasia in adults and SLI in children which are characterized by various profiles of language impairment may be remediated

by a similar intervention program based on non-verbal training in TIP. Specifically, post-stroke receptive aphasia investigated in Study 1 is characterized by relatively fluent verbal output but disordered auditory comprehension. In contrast, SLI investigated in Study 2 is characterized by developmental language production and/or comprehension deficits (usually mixed) that cannot be explained by general cognitive impairment, concomitant impairments or a general lack of exposure to language. Whereas the etiology of aphasia is usually well defined and the lesioned area may be evidenced in neuroimaging examination (see **Figure 1**), the etiology of SLI is difficult to define and remains predominantly unknown.

The exciting phenomenon of human language is that, despite the totally different pattern of language impairment in case of these two disorders, there are some important similarities in neuronal mechanisms underlying disordered language. These mechanisms are rooted in temporal acuity on millisecond range and can be studied using both verbal (temporal dynamic of the spontaneous speech) or non-verbal information processing (indexed by ATOT). The important finding is that these distinct language impairments at least in some cases are sensitive to the specific training focused on TIP. Such finding can help to design and elaborate future remediation programs supporting the classic speech therapy.

A final question is who more could benefit from such therapy program. In our previous studies (Szelag and Skolimowska, 2012) we indicated that the temporal frame does not underlie selectively speech processing, but also some other cognitive functions, like working memory, attention or executive control which can be characterized also by the specific temporal dynamics. Moreover, these cognitive functions could be also remediated following the specific training (**Figures 3 and 4**, see the Results). We conclude, therefore, that the future

horizons for the application of the temporal intervention may be expanded by applications in enhancement of the broad aspects of cognitive functioning. Such view point may be supported by the results of our previous study in which we indicated that the application of FFW training in normal healthy elderly beyond 65 years of life resulted in improved attention and short-term memory (Szelag and Skolimowska, 2012).

To sum up, amelioration of disordered timing can be used as an universal tool in future clinical practice not only in language-disordered population, but also in people with various cognitive dysfunctions.

## FUNDING

The research was supported by grant INNOTECH-K1/IN1/30/159041/NCBR/12 (The National Centre for Research and Development, Poland).

## ACKNOWLEDGMENTS

The authors would like to thank neuropsychologists and speech-language pathologists who recruited aphasic patients and children with SLI for the present studies. We thank Elzbieta Chruscicka and Monika Kastory-Bronowska<sup>†</sup> from the Early Intervention Centre in Warsaw for their assistance in recruitment of children into Study 2. The prototyping training applied in Study 1 and the computer program for ATOT assessment were elaborated in cooperation with Marc Wittmann, Martina Fink, Pamela Ulbrich and Jan Churan from the Human Science Centre, Ludwig-Maximilian University, Bad Tölz, Germany.

## REFERENCES

- Alarcón-Rubio, D., Sánchez-Medina, J. A., and Prieto-García, J. R. (2014). Executive function and verbal self-regulation in childhood: developmental linkages between partially internalized private speech and cognitive flexibility. *Early Child. Res. Q.* 29, 95–105. doi: 10.1016/j.jecresq.2013.11.002
- Baddeley, A. (2003). Working memory and language: an overview. *J. Commun. Disord.* 36, 189–208. doi: 10.1016/S0021-9924(03)00019-4
- Bailey, P. J., and Snowling, M. J. (2002). Auditory processing and the development of language and literacy. *Br. Med. Bull.* 63, 135–146. doi: 10.1093/bmb/63.1.135
- Bao, Y., Szymaszek, A., Wang, X., Oron, A., Pöppel, E., and Szelag, E. (2013). Temporal order perception of auditory stimuli is selectively modified by tonal and non-tonal language environments. *Cognition* 129, 579–585. doi: 10.1016/j.cognition.2013.08.019
- Bavelier, D., Green, C. S., Han, D. H., Renshaw, P. F., Merzenich, M. M., and Gentile, D. A. (2011). Brains on video games. *Nat. Rev. Neurosci.* 12, 763–768. doi: 10.1038/nrn3135
- Cambridge Cognition (2005). *Neuropsychological Test Automated Battery (CANTABclipse) Manual*. Cambridge: Cambridge Cognition Limited.
- Cardy, J. E. O., Tannock, R., Johnson, A. M., and Johnson, C. J. (2010). The contribution of processing impairments to SLI: insights from attention-deficit/hyperactivity disorder. *J. Commun. Disord.* 43, 77–91. doi: 10.1016/j.jcomdis.2009.09.003
- Cohen, W., Hodson, A., O'Hare, A., Boyle, J., Durrani, T., McCartney, E., et al. (2005). Effects of computer-based intervention through acoustically modified speech (Fast ForWord) in severe mixed receptive - expressive language impairment outcomes from a randomized controlled trial. *J. Speech Lang. Hear. Res.* 48, 715–729. doi: 10.1044/1092-4388(2005/049)
- Culbertson, W. C., and Zillmer, E. A. (2005). *Tower of London – Drexel University (TOLDX), Technical Manual*, 2nd Edn. Toronto: Multi-Health Systems.
- Efron, R. (1963). Temporal perception, aphasia and déjà vu. *Brain* 86, 403–424. doi: 10.1093/brain/86.3.403
- Fink, M., Churan, J., and Wittmann, M. (2006). Temporal processing and context dependency of phoneme discrimination in patients with aphasia. *Brain Lang.* 98, 1–11. doi: 10.1016/j.bandl.2005.12.005
- Gaab, N., Gabrieli, J. D. E., Deutsch, G. K., Tallal, P., and Temple, E. (2007). Neural correlates of rapid auditory processing are disrupted in children with developmental dyslexia and ameliorated with training: an fMRI study. *Restor. Neurol. Neurosci.* 25, 295–310.
- Gathercole, S. E., and Baddeley, A. D. (1990). Phonological memory deficits in language disordered children: is there a causal connection? *J. Mem. Lang.* 29, 336–360. doi: 10.1016/0749-596X(90)90004-J
- Gillam, R. B., Loeb, D. F., Hoffman, L. M., Bohman, T., Champlin, C. A., Thibodeau, L., et al. (2008). The efficacy of Fast ForWord language intervention in school-age children with language impairment: a randomized controlled trial. *J. Speech Lang. Hear. Res.* 51, 97–119. doi: 10.1044/1092-4388(2008/007)

- Given, B. K., Wasserman, J. D., Chari, S. A., Beattie, K., and Eden, G. F. (2008). A randomized, controlled study of computer-based intervention in middle school struggling readers. *Brain Lang.* 106, 83–97. doi: 10.1016/j.bandl.2007.12.001
- Grondin, S., Dionne, G., Malenfant, N., Plourde, M., Cloutier, M. E., and Jean, C. (2007). Temporal processing skills of children with and without specific language impairment. *Can. J. Speech Lang. Pathol. Audiol.* 31, 38–46.
- Hayes, E. A., Warrier, C. M., Nicol, T. G., Zecker, S. G., and Kraus, N. (2003). Neural plasticity following auditory training in children with learning problems. *Clin. Neurophysiol.* 114, 673–684. doi: 10.1016/S1388-2457(02)00414-5
- Heiervang, E., Stevenson, J., and Hugdahl, K. (2002). Auditory processing in children with dyslexia. *J. Child Psychol. Psychiatry* 43, 931–938. doi: 10.1111/1469-7610.00097
- Heim, S., Keil, A., Choudhury, N., Friedman, J. T., and Benasich, A. A. (2013). Early gamma oscillations during rapid auditory processing in children with a language-learning impairment: changes in neural mass activity after training. *Neuropsychologia* 51, 990–1001. doi: 10.1016/j.neuropsychologia.2013.01.011
- Henry, L. A., Messer, D. J., and Nash, G. (2012). Executive functioning in children with specific language impairment. *J. Child Psychol. Psychiatry* 53, 37–45. doi: 10.1111/j.1469-7610.2011.02430.x
- Huber, W., Poeck, K., Weniger, D., and Willmes, K. (1983). *Aachener Aphasia Test (AAT)*. Göttingen: Hogrefe.
- Koscieszka, M., and Krasowicz, G. (1995). Polska adaptacja techniki “Token Test” dla dzieci. Możliwości praktycznego zastosowania. *Psychol. Wychowawcza* 4, 350–359.
- Lewandowska, M., Piatkowska-Janko, E., Bogorodzki, P., Wolak, T., and Szelag, E. (2010). Changes in fMRI BOLD response to increasing and decreasing task difficulty during auditory perception of temporal order. *Neurobiol. Learn. Mem.* 94, 382–391. doi: 10.1016/j.nlm.2010.08.005
- Lezak, M. D. (1995). *Neuropsychological Assessment*. New York, NY: Oxford University Press.
- Lucki, W. (1995). *Zestaw Prob do Badania Procesow Poznawczych u Pacjentow z Uszkodzeniami Mozgu*. Warszawa: Pracownia Testow Psychologicznych.
- Matczak, A., Piotrowska, A., and Ciarkowska, W. (1991). *Skala Inteligencji D. Wechslera dla Dzieci – Wersja Zmodyfikowana (WISC-R)*. Podrecznik. Warszawa: Polskie Towarzystwo Psychologiczne.
- Nowak-Czerwinska, M. (1994). “Test sluchowego roznicowania glosek (TSRG),” in *Mowa Glosna i pismo. Materiały z Konferencji Naukowej Sekcji Logopedycznej Towarzystwa Kultury Języka* (Białystok: Towarzystwo Kultury Języka), 10–13.
- Oron, A., Szymaszek, A., and Szelag, E. (2015). Temporal information processing as a basis for auditory comprehension: clinical evidence from aphasic patients. *Int. J. Lang. Comm. Dis.* 50, 604–615. doi: 10.1111/1460-6984.12160
- Poeppel, D. (2003). The analysis of speech in different temporal integration windows: cerebral lateralization as ‘asymmetric sampling in time.’ *Speech Commun.* 41, 245–255. doi: 10.1016/S0167-6393(02)00107-3
- Pöppel, E. (1994). Temporal mechanisms in perception. *Int. Rev. Neurobiol.* 37, 185–202. doi: 10.1016/S0074-7742(08)60246-9
- Pöppel, E. (1997). A hierarchical model of temporal perception. *Trends Cogn. Sci.* 1, 56–61. doi: 10.1016/S1364-6613(97)01008-5
- Pöppel, E. (2009). Pre-semantically defined temporal windows for cognitive processing. *Philos. Trans. R. Soc. B Biol. Sci.* 364, 1887–1896. doi: 10.1098/rstb.2009.0015
- Ramus, F., Rosen, S., Dakin, S. C., Day, B. L., Castellote, J. M., White, S., et al. (2003). Theories of developmental dyslexia: insights from a multiple case study of dyslexic adults. *Brain* 126, 841–865. doi: 10.1093/brain/awg076
- Rice, M. L., and Wexler, K. (1996). Toward tense as a clinical marker of specific language impairment in English-speaking children. *J. Speech Hear. Res.* 39, 1239–1257. doi: 10.1044/jshr.3906.1239
- Roello, M., Ferretti, M. L., Colonnello, V., and Levi, G. (2015). When words lead to solutions: executive function deficits in preschool children with specific language impairment. *Res. Dev. Disabil.* 37, 216–222. doi: 10.1016/j.ridd.2014.11.017
- Share, D. L., Jorm, A. F., Maclean, R. O. D., and Matthews, R. (2002). Temporal processing and reading disability. *Read. Writ.* 15, 151–178. doi: 10.1023/A:1013876606178
- Sidiropoulos, K., Ackermann, H., Wannke, M., and Hertrich, I. (2010). Temporal processing capabilities in repetition conduction aphasia. *Brain Cogn.* 73, 194–202. doi: 10.1016/j.bandc.2010.05.003
- Stefanatos, G. A., Braitman, L. E., and Madigan, S. (2007). Fine grain temporal analysis in aphasia: evidence from auditory gap detection. *Neuropsychologia* 45, 1127–1133. doi: 10.1016/j.neuropsychologia.2006.09.011
- Stevens, C., Fanning, J., Coch, D., Sanders, L., and Neville, H. (2008). Neural mechanisms of selective auditory attention are enhanced by computerized training: electrophysiological evidence from language-impaired and typically developing children. *Brain Res.* 1205, 55–69. doi: 10.1016/j.brainres.2007.10.108
- Swisher, L., and Hirsh, I. J. (1972). Brain damage and the ordering of two temporally successive stimuli. *Neuropsychologia* 10, 137–152. doi: 10.1016/0028-3932(72)90053-X
- Szelag, E., Dreszer, J., Lewandowska, M., and Szymaszek, A. (2009). “Cortical representation of time and timing processes,” in *Neuronal Correlates of Thinking*, eds E. Kraft, B. Guylas, and E. Pöppel (Berlin: Springer-Verlag), 185–199.
- Szelag, E., Kanabus, M., Kolodziejczyk, I., Kowalska, J., and Szuchnik, J. (2004). Individual differences in temporal information processing in humans. *Acta Neurobiol. Exp.* 64, 349–366.
- Szelag, E., Lewandowska, M., Wolak, T., Seniow, J., Poniatowska, R., Pöppel, E., et al. (2014). Training in rapid auditory processing ameliorates auditory comprehension in aphasic patients: a randomized controlled pilot study. *J. Neurol. Sci.* 338, 77–86. doi: 10.1016/j.jns.2013.12.020
- Szelag, E., Skarzynski, H., Senderski, A., and Lewandowska, M. (2011a). “Hearing loss and auditory processing disorders: clinical and experimental perspectives,” in *Culture and Neural Frames of Cognition and Communication*, eds S. Han and E. Pöppel (Berlin: Springer), 153–168.
- Szelag, E., Szymaszek, A., Aksamit-Ramotowska, A., Fink, M., Ulbrich, P., Wittmann, M., et al. (2011b). Temporal processing as a base for language universals: cross-linguistic comparisons on sequencing abilities with some implications for language therapy. *Restor. Neurol. Neurosci.* 29, 35–45. doi: 10.3233/RNN-2011-0574
- Szelag, E., and Skolimowska, J. (2012). Cognitive function in elderly can be ameliorated by training in temporal information processing. *Restor. Neurol. Neurosci.* 30, 419–434. doi: 10.3233/RNN-2012-120240
- Szelag, E., and Szymaszek, A. (2006). *Test do Badania Sluchu Fonematycznego u Dzieci i Doroslych*. Gdansk: Gdanskie Wydawnictwo Psychologiczne.
- Szelag, E., and Szymaszek, A. (2014). *Test do Badania Rozumienia Mowy u Dzieci i Doroslych*. Gdansk: Gdanskie Wydawnictwo Psychologiczne.
- Szelag, E., Szymaszek, A., and Oron, A. (2015). “Aphasia as temporal information processing disorder,” in *Time Distortions in Mind – Temporal Processing in Clinical Populations*, eds A. Vatakis and M. J. Allman (Boston, MA: Brill Publications), 328–355.
- Szepietowska, E. M. (2000). *Badanie Neuropsychologiczne. Procedura i Ocena*. Lublin: Wydawnictwo UMCS.
- Szustrowa, T., and Jaworowska, A. (2003). *Manual for Raven’s Progressive Matrices. Colored Progressive Matrices*. Warszawa: Polskie Towarzystwo Psychologiczne.
- Szymaszek, A., Sereda, M., Pöppel, E., and Szelag, E. (2009). Individual differences in the perception of temporal order: the effect of age and cognition. *Cogn. Neuropsychol.* 26, 135–147. doi: 10.1080/02643290802504742
- Tallal, P., Miller, S. L., Bedi, G., Byma, G., Wang, X., Nagarajan, S. S., et al. (1996). Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science* 271, 81–84. doi: 10.1126/science.271.5245.81
- Tallal, P., and Piercy, M. (1973). Deficits of non-verbal auditory perception in children with developmental aphasia. *Nature* 241, 468–469. doi: 10.1038/241468a0
- Tallal, P., and Piercy, M. (1974). Developmental aphasia: rate of auditory processing and selective impairment of consonant perception. *Neuropsychologia* 12, 83–93. doi: 10.1016/0028-3932(74)90030-X
- Tarkowski, Z. (2001). *Test Sprawnosci Językowej*. Lublin: Orator.



- Temple, E., Deutsch, G. K., Poldrack, R. A., Miller, S. L., Tallal, P., Merzenich, M. M., et al. (2003). Neural deficits in children with dyslexia ameliorated by behavioral remediation: evidence from functional MRI. *Proc. Natl. Acad. Sci. U.S.A.* 100, 2860–2865. doi: 10.1073/pnas.0030098100
- Treutwein, B. (1997). YAAP: yet another adaptive procedure. *Spat. Vis.* 11, 129–134.
- Ulbrich, P., Churan, J., Fink, M., and Wittmann, M. (2009). Perception of temporal order: the effect of age, sex and cognitive factors. *Neuropsychol. Dev. Cogn. B Aging Neuropsychol. Cogn.* 16, 183–202. doi: 10.1080/13825580802411758
- Ullman, M. T., and Pierpont, E. I. (2005). Specific language impairment is not specific to language: the procedural deficit hypothesis. *Cortex* 41, 399–433. doi: 10.1016/S0010-9452(08)70276-4
- van Rullen, R., and Koch, C. (2003). Is perception discrete or continuous? *Trends Cogn. Sci.* 7, 207–213. doi: 10.1016/S1364-6613(03)00095-0
- van Zomeren, A. H., and Bouwer, W. H. (1994). *Clinical Neuropsychology of Attention*. New York, NY: Oxford University Press.
- Wittmann, M., Bartscher, A., Fries, W., and von Steinbüchel, N. (2004). Effects of brain-lesion size and location on temporal-order judgment. *Neuroreport* 15, 2401–2405. doi: 10.1097/00001756-200410250-00020
- World Health Organization (1992). *The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines*. Geneva: World Health Organization.
- Zimmermann, P., and Fimm, B. (1997). *Test of Attentional Performance*. Herzogenrath: Psytest.
- Zimmermann, P., Gonda, M., and Fimm, B. (2005). *Test of Attentional Performance for Children*. Herzogenrath: Psytest.
- Conflict of Interest Statement:** 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.

Copyright © 2015 Szelag, Dacewicz, Szymaszek, Wolak, Senderski, Domitrz and Oron. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Corrigendum: The Application of Timing in Therapy of Children and Adults with Language Disorders

Elzbieta Szelag<sup>1,2\*</sup>, Anna Dacewicz<sup>1</sup>, Aneta Szymaszek<sup>1,2</sup>, Tomasz Wolak<sup>3</sup>, Andrzej Senderski<sup>4</sup>, Izabela Domitrz<sup>5</sup> and Anna Oron<sup>1</sup>

<sup>1</sup> Laboratory of Neuropsychology, Nencki Institute of Experimental Biology, Warsaw, Poland, <sup>2</sup> University of Social Sciences and Humanities, Warsaw, Poland, <sup>3</sup> Institute of Physiology and Pathology of Hearing, Kajetany, Poland, <sup>4</sup> Children's Memorial Health Institute, Warsaw, Poland, <sup>5</sup> Department of Neurology, Warsaw Medical University, Warsaw, Poland

**Keywords:** temporal information processing, language, cognitive functions, aphasia, specific language disorder

## A corrigendum on

### The Application of Timing in Therapy of Children and Adults with Language Disorders

by Szelag, E., Dacewicz, A., Szymaszek, A., Wolak, T., Senderski, A., Domitrz, I., et al. (2015). *Front. Psychol.* 6:1714. doi: 10.3389/fpsyg.2015.01714

## OPEN ACCESS

### Edited and reviewed by:

Lihan Chen,  
Peking University, China

### \*Correspondence:

Elzbieta Szelag  
e.szelag@nencki.gov.pl

### Specialty section:

This article was submitted to  
Perception Science,  
a section of the journal  
*Frontiers in Psychology*

**Received:** 09 March 2016

**Accepted:** 11 March 2016

**Published:** 30 March 2016

### Citation:

Szelag E, Dacewicz A, Szymaszek A, Wolak T, Senderski A, Domitrz I and Oron A (2016) Corrigendum: The Application of Timing in Therapy of Children and Adults with Language Disorders. *Front. Psychol.* 7:449. doi: 10.3389/fpsyg.2016.00449

## Conflict of Interest Statement

ES and AS are the creators of the software package Dr. Neuronowski®, realized as part of a project at the Nencki Institute with funding from the National Centre for Research and Development in Poland. The rights to the software lie with the Nencki Institute that has an agreement with Harpo Ltd., the company commercializing this software. ES and AS are not the owners of this technology nor do they have a direct financial arrangement with Harpo Ltd.

The authors state that the correction does not affect the scientific validity of the results.

## AUTHOR CONTRIBUTIONS

ES: experimental design, data acquisition and analysis, manuscript writing. AD: data acquisition and analysis, manuscript writing. AS: data acquisition and analysis, manuscript writing. TW: guidance on MRI data analysis. AS: subject recruitment. ID: subject recruitment. AO: data acquisition and analysis, manuscript writing.

**Conflict of Interest Statement:** The other 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.

Copyright © 2016 Szelag, Dacewicz, Szymaszek, Wolak, Senderski, Domitrz and Oron. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Early, but not late visual distractors affect movement synchronization to a temporal-spatial visual cue

Ashley J. Booth<sup>1</sup> and Mark T. Elliott<sup>1,2\*</sup>

<sup>1</sup> School of Psychology, University of Birmingham, Edgbaston, UK, <sup>2</sup> Institute of Digital Healthcare, Warwick Manufacturing Group, University of Warwick, Coventry, UK

## OPEN ACCESS

### Edited by:

Lihan Chen,  
Peking University, China

### Reviewed by:

Yoshimori Sugano,  
Kyushu Sangyo University, Japan  
Yi-Huang Su,  
Technical University of Munich,  
Germany

### \*Correspondence:

Mark T. Elliott,  
Institute of Digital Healthcare,  
Warwick Manufacturing Group,  
University of Warwick, University  
Road, Coventry CV4 7AL, UK  
m.t.elliott@warwick.ac.uk

### Specialty section:

This article was submitted to  
Perception Science,  
a section of the journal  
Frontiers in Psychology

**Received:** 20 April 2015

**Accepted:** 12 June 2015

**Published:** 24 June 2015

### Citation:

Booth AJ and Elliott MT (2015) Early,  
but not late visual distractors affect  
movement synchronization  
to a temporal-spatial visual cue.  
Front. Psychol. 6:866.  
doi: 10.3389/fpsyg.2015.00866

The ease of synchronizing movements to a rhythmic cue is dependent on the modality of the cue presentation: timing accuracy is much higher when synchronizing with discrete auditory rhythms than an equivalent visual stimulus presented through flashes. However, timing accuracy is improved if the visual cue presents spatial as well as temporal information (e.g., a dot following an oscillatory trajectory). Similarly, when synchronizing with an auditory target metronome in the presence of a second visual distracting metronome, the distraction is stronger when the visual cue contains spatial-temporal information rather than temporal only. The present study investigates individuals' ability to synchronize movements to a temporal-spatial visual cue in the presence of same-modality temporal-spatial distractors. Moreover, we investigated how increasing the number of distractor stimuli impacted on maintaining synchrony with the target cue. Participants made oscillatory vertical arm movements in time with a vertically oscillating white target dot centered on a large projection screen. The target dot was surrounded by 2, 8, or 14 distractor dots, which had an identical trajectory to the target but at a phase lead or lag of 0, 100, or 200 ms. We found participants' timing performance was only affected in the phase-lead conditions and when there were large numbers of distractors present (8 and 14). This asymmetry suggests participants still rely on salient events in the stimulus trajectory to synchronize movements. Subsequently, distractions occurring in the window of attention surrounding those events have the maximum impact on timing performance.

**Keywords:** sensorimotor synchronization, visual cues, movement timing, distractor cues

## Introduction

Nodding or tapping along to a favorite song is often something we do with little conscious thought. This demonstrates the automaticity of being able to move in time to a rhythmic stimulus, an ability that forms the basis of sensorimotor synchronization (SMS) research (Repp and Su, 2013). The majority of SMS research has focussed on the timing of movements to an auditory rhythmic cue and indeed it appears this is the sensory modality that facilitates the most accurate timing of movements (Repp and Penel, 2004; Elliott et al., 2010). However, movement synchrony can also occur outside the context of music. In social situations, groups of individuals can spontaneously coordinate the timing of their movements, for example, two people falling into step when walking together (Zivotofsky et al., 2012), or an excited crowd bouncing up and down together in a sports stadium (Noormohammadi et al., 2011). In these group scenarios, visual cues are likely to

provide a strong timing stimulus that results in implicit synchrony emerging within the group. However, with each person in a group exhibiting slightly different timing properties, it is currently unclear how synchrony occurs in the face of conflicting visual cues. Here, we have developed an experimental paradigm that investigates how individuals synchronize movements to a target visual cue in the presence of conflicting visual stimuli.

Timing accuracy in SMS studies is often quantified by the *asynchronies*, which represent the time difference between the target and the executed movement. The mean and variability of the asynchronies are taken into account. A negative mean asynchrony (NMA) is usually observed in SMS research where the movement typically precedes the target by 30–50 ms (Aschersleben and Prinz, 1995). While auditory cues dominate SMS research, other modalities have been investigated. In particular, SMS to a discrete flashing visual stimulus results in reduced timing accuracy in terms of asynchrony variability (Repp and Penel, 2004; Kurgansky, 2008; Elliott et al., 2010; Wright and Elliott, 2014) compared to an auditory metronome. Hence, discrete auditory stimuli provide a more reliable, salient cue compared to a discrete rhythmic visual cue (Repp and Penel, 2004). However, more recent studies found that synchronizing movement to continuous visual cues, i.e., those exhibiting temporal and spatial information, yielded strong SMS that was comparable to studies using auditory cues (Hove et al., 2012; Varlet et al., 2012; Armstrong et al., 2013). Moreover, visual trajectories representing biologically compatible movements further facilitates rhythm perception (Su, 2014a) and movement synchronization (Su, 2014c). This latter finding indicates how the temporal-spatial visual information provided by surrounding members of a group could influence the implicit synchrony of movements within the group.

A number of studies have implemented a distractor paradigm to observe how irrelevant cues presented in auditory or auditory versus visual modalities can affect an individual's ability to synchronize their movements to a target cue. As might be expected, an auditory distractor in the presence of a discrete visual target leads to a strong distraction effect, due to the strong saliency of the auditory modality (Repp and Penel, 2002, 2004). These distraction effects are quantified through a change in NMA, i.e., asynchronies becoming more negative in the presence of early distractors or more positive for late distractors, and asynchrony variability, with strong distractor effects reducing the stability of the asynchronies. In general, discrete distractor cues (be it auditory-auditory or auditory-visual modalities) exhibit an asymmetric NMA effect, where a strong attraction is observed when the distractor precedes the target, but show little change for late distractors (Repp, 2003; Repp and Penel, 2004).

What is currently unclear is how an individual's ability to synchronize movements to temporal-spatial visual cues is affected by similar conflicting visual distractors. In this study, we investigated participants' ability to synchronize oscillatory arm movements in time to a temporal-spatial oscillating visual target, in the presence of identical visual distractors offset in phase to the target. As well as varying phase to influence the temporal relation between target and distractor, we also varied the visual impact of the distraction effect by varying the number of distractor

stimuli present. Increasing the number of distraction stimuli should correspondingly increase visual attention to the distractors (Bartram et al., 2003). Hence we predicted that the strength of the distraction effect would be a function of both the temporal separation and the number of distractors present. As observed in previous studies, we further expected that the temporal distraction would be at its greatest when the phase offset was around a quarter of the oscillation period (Repp, 2003; Repp and Penel, 2004). However, due to the continuous nature of both the movements and the stimuli, we did not expect to see an asymmetry in the distraction effect as observed with discrete stimuli paradigms.

## Materials and Methods

### Participants

Eleven University of Birmingham undergraduate Psychology students (six female;  $M_{age} = 18.4$ , range = 18–20, SD 0.67 years) gave written informed consent to take part in the study. All participants reported themselves free of any neurological disease, head trauma, musculoskeletal impairment, visual impairment, or hearing impairment. Ethical approval was granted by the University of Birmingham Science, Technology, Engineering, and Mathematics Ethical Review Committee. Of the 11 participants, nine were right-handed. Data from one participant was removed due to difficulty with following instructions and completing the task correctly.

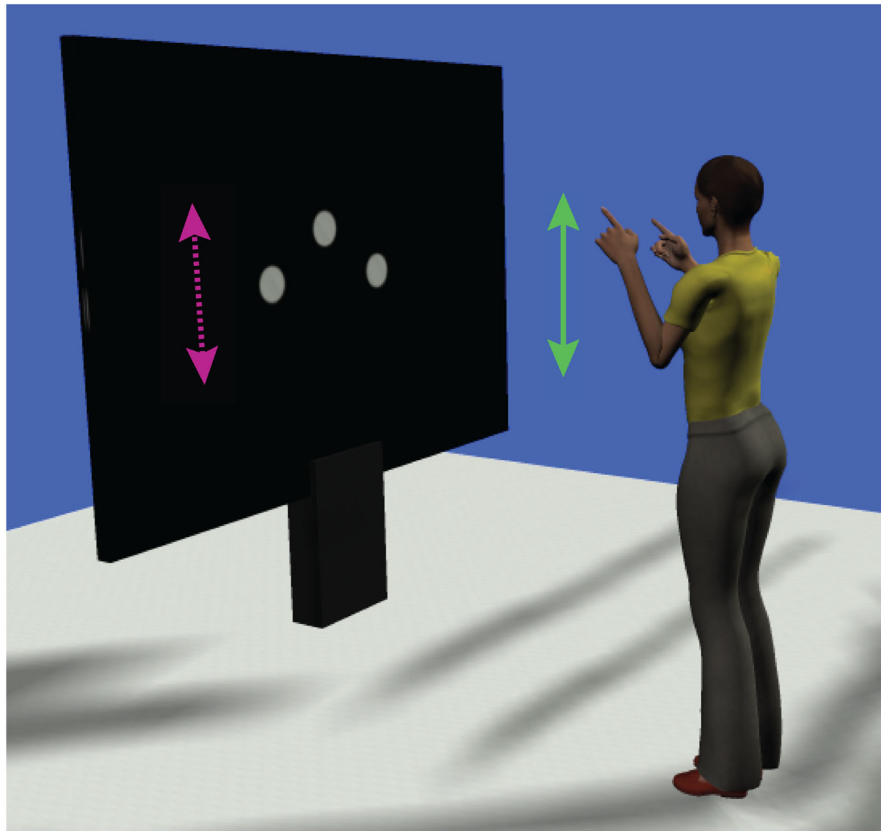
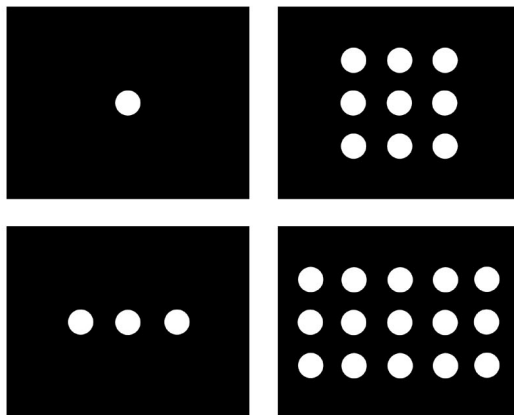
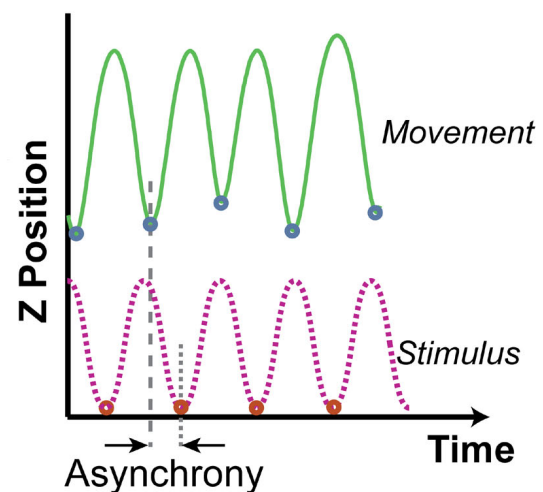
### Experimental Setup

Participants stood on a marked point 1.85 m from a projection screen (1.6 m wide  $\times$  1.2 m tall; **Figure 1A**). Arm movement trajectories were captured using a 12-camera Qualisys Oqus motion capture system (Qualisys AB, Gothenburg, Sweden), with adhesive reflective markers attached to the shoulders, elbows, wrists, and index fingers of both arms. The camera system operated with a sampling rate of 200 Hz.

### Stimuli

Visual stimuli were generated in Matlab (2013a; The Mathworks, MA, USA) Psychophysics Toolbox (Brainard, 1997). The stimuli consisted of a series of white circular dots (100 pixels diameter) moving vertically against a black background with a sinusoidal trajectory (period: 800 ms, 60 frames per second). The peak-peak range of movement for the dots was 200 pixels. Participants were instructed to synchronize movements with the “target”—a centrally positioned dot that was present in all conditions. In addition, a number of distractor dots were positioned symmetrically to the sides, above and below the target. There were four distractor conditions, which consisted of 0, 2, 8, or 14 distractor dots in the formations shown in **Figure 1B**. Dots were separated from one another, center to center, by 125 pixels horizontally and 200 pixels vertically. In addition to the different numbers of distractors, there were five “phase-offset” conditions where the timing of the distractor dots was offset such there was a constant phase lead (negative) or lag (positive) of 0,  $\pm 100$ , or  $\pm 200$  ms relative to the central target trajectory. The spacing of the dots was designed such that none of the phase-offset conditions resulted in occlusion of the target dot on the screen. A digital high



**A****B****C**

**FIGURE 1 | (A)** Representation of experimental set up. Participants faced a large projection screen which presented the visual stimuli. The stimuli (100 pixel diameter dots) moved vertically up and down, following a sinusoidal trajectory. The target stimulus was always the center dot. Distractor dots moved out of phase with the target by 0,  $\pm 100$ ,  $\pm 200$  ms. Participants made bimanual arm movements in synchrony with the target stimulus, flexing and extending the forearm from the elbow. **(B)** Formation of target and distractor stimuli. We investigated if the distraction effect was a function of the number of distractor stimuli. The number of distractor stimuli was

varied across trials such that there were no distractors (top left), two distractors (bottom left), eight distractors (top right), or 14 distractors (bottom right). **(C)** Measurements of timing accuracy. Representative trajectories of the target stimulus (bottom trace, dashed pink) and the corresponding participant's dominant arm movement (top, solid green) are shown. We extracted the times of the minimum positions for each movement oscillation along with the times of the minimum stimulus positions. Asynchrony was calculated by subtracting the time of movement event from the time of the corresponding stimulus event.

(+5V) signal pulse was output via a data acquisition card (USB-6343, National Instruments, TX, USA) to the Qualisys motion capture system each time the target dot reached its minimum position in the trajectory. This was used to align screen output with the participant's movements (see Data Processing).

## Experimental Design and Procedure

Participants completed the study individually. They were instructed to move both forearms up and down in synchrony with the central target dot, flexing and extending at the elbows with only their index fingers extended. We instructed the use of bimanual movements to improve timing stability (Helmuth and Ivry, 1996). In addition, during pilot tests participants reported bimanual movements to be more comfortable and natural for the task. Participants were further required to keep their wrists tense and so were instructed to keep their wrists firm such that a straight line could be imagined between the fingertip and elbow during the movement. They were told to ignore the movements of the non-target dots to the best of their ability. A practice trial was carried out to ensure that the requirements were fully understood and they were ready to continue.

There were three trials for each condition (3 Distractors conditions: 2, 8, 14  $\times$  5 Phase offset conditions: -200, -100, 0, 100, 200 ms; plus a no-distractor condition) totalling 48 trials in all. The order of the trials was randomized for each participant to avoid order effects. Each trial lasted 40 s, which resulted in 50 dot oscillations per trial.

## Data Processing

Only the vertical (z-axis) data from the reflective marker attached to the index finger on the dominant hand was used for analysis. Using a peak detection algorithm from the MatTAP toolbox (Elliott et al., 2009b), the "event times" of the lowest vertical points of the executed oscillatory arm movements were extracted (Figure 1C). Lowest points were chosen as evidence suggests synchronization is more stable on the downward movement (Miura et al., 2011). Similarly, the event times of the lowest positions of the target stimulus were recorded as the time at which the digital signal from data acquisition was set high (see Stimuli). The first five event times from each trial were discarded from the analysis to allow for participants to initially synchronize with the target. The event times between the stimulus and the participant's movements were then aligned (Elliott et al., 2009b) by finding the movement onset time closest to each stimulus onset time (on average <1% of all stimulus onsets could not be aligned to a participant's corresponding movement, indicating participants were able to perform the task). Subsequently, the asynchronies were calculated as the time difference between the stimulus event and the corresponding movement event. A negative asynchrony indicated that the movement event occurred before the stimulus (Figure 1C).

The standard deviation and mean asynchrony were calculated for each trial and the average taken across trials for each participant. We initially analyzed the effects of the number of distractors and phase offset (reported in sections "Mean Asynchrony" and "Standard Deviation") using a 3 (Distractors: 2, 8, 14)  $\times$  5 (Phase Offset: -200, -100, 0, 100, 200 ms)

repeated measures design. We further analyzed just the effect of number of Distractors using data from the 0 phase-offset conditions in addition to the baseline "no distractor" condition [4 (Distractors: 0, 2, 8, 14)  $\times$  1 (Phase Offset: 0 ms) repeated measures; reported in section "Comparison of No-Distractor with Distractor Conditions"]. Statistical analysis was completed using Repeated Measures ANOVAs in SPSS (version 21, IBM Corp., NY, USA). Significance levels were set to  $p < 0.05$ . Greenhouse-Geisser adjustments were made for results that violated sphericity assumptions. *Post hoc* analyses were adjusted for multiple comparisons using the Bonferroni method.

## Results

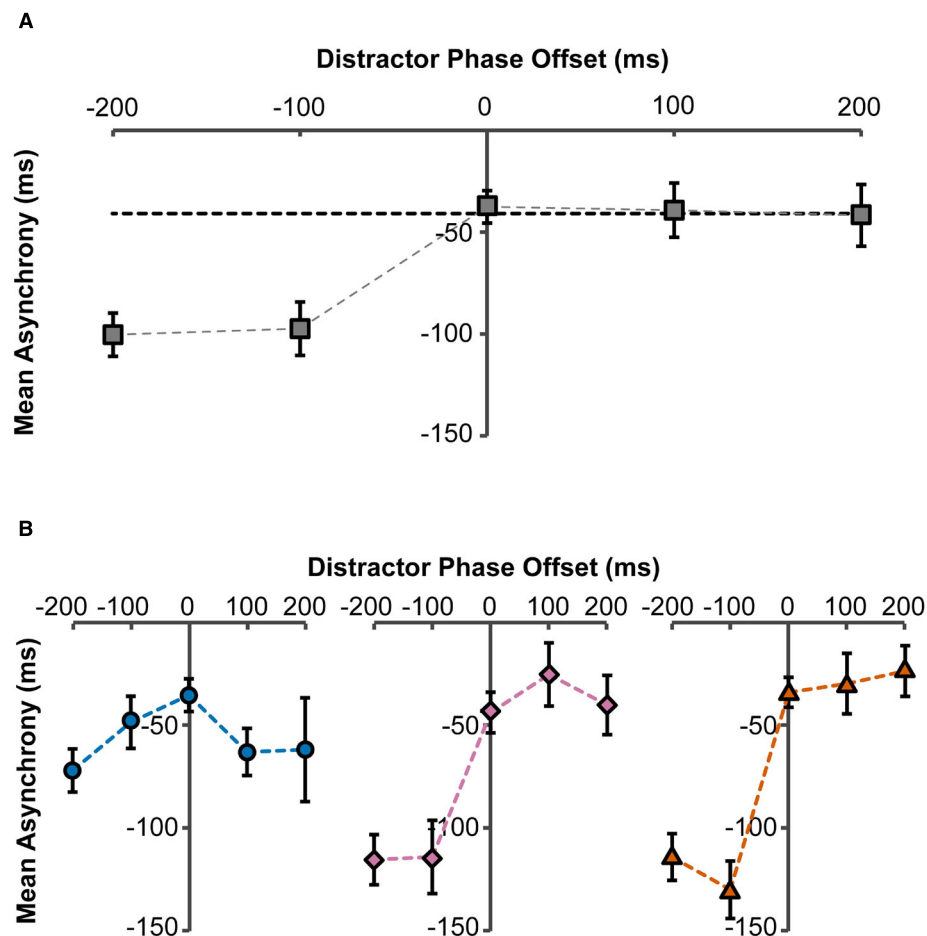
### Mean Asynchrony

A repeated measures within-participants ANOVA revealed that there was a significant effect of phase-offset on mean asynchrony [ $F(4,36) = 25.17, p < 0.001$ ]. That is, changes to the phase-offset significantly affected synchronization to the target (Figure 2A). *Post hoc* analysis identified that there were only significant differences between the 0 ms phase-offset condition relative to the -200 ms condition ( $M = -62.8$  ms,  $p < 0.001$ ) and the -100 ms condition ( $M = -59.8$  ms,  $p = 0.001$ ). However, there were no significant differences between the -200 ms and -100 ms phase-offsets conditions, and so performance does not continue to decline linearly as the phase-offset increases. Moreover, the positive phase offsets did not significantly alter the mean asynchrony compared to the 0 ms phase-offset. These findings show that there is an asymmetrical effect of phase-offset where the negative phase-offset conditions make the mean asynchrony more negative, so arm movements were drawn to the phase-leading distractor trajectories. In contrast, movements were not drawn to phase-lagging distractor trajectories.

There was no significant main effect of the number of distractors on the mean asynchrony; however, the analysis yielded a significant interaction between the number of distractors and phase-offset [ $F(2.7,24.4) = 13.36, p < 0.001$ ; Figure 2B]. Analyzing each Distractor condition separately highlighted that when only two distractors were present, there was no effect of phase-offset on the mean asynchrony [ $F(1.58,14.26) = 1.83, p = 0.199$ ]. In contrast, for the 8 dot [ $F(4,36) = 32.79, p < 0.001$ ] and 14 dot [ $F(4,36) = 36.48, p < 0.001$ ] distractor conditions, the previously described phase attraction for leading distractors was present (Figure 2B).

### Standard Deviation

We further investigated how the distractors impacted on the variability (standard deviation) of the asynchronies over a trial. Again, we observed a significant main effect of phase-offset [ $F(4,36) = 5.14, p = 0.002$ ; Figure 3]. *Post hoc* analyses identified the -100 ms phase-offset as the only condition that significantly differed from the 0 ms phase-offset condition ( $M$  difference = 10.5,  $p = 0.014$ ). We found that in this condition, the variability of asynchronies significantly increased, indicating that the strongest distraction occurred when the distractor stimuli were moving earlier in phase by around 100 ms.



**FIGURE 2 | Mean asynchrony as a function of distractor phase-offset to target stimulus.** Asynchronies were measured between the participant's movements and the target stimulus. **(A)** Overall effect of distractor phase-offset on mean asynchrony, collapsed across number of distractors. Error bars represent standard errors. Dashed

horizontal black line indicates baseline mean asynchrony in the no-distractor condition. **(B)** Effect of distractor phase-offset on mean asynchrony, with two distractors present (circles), eight distractors present (diamonds), and 14 distractors present (triangles). Error bars represent standard errors.

## Comparison of No-Distractor with Distractor Conditions

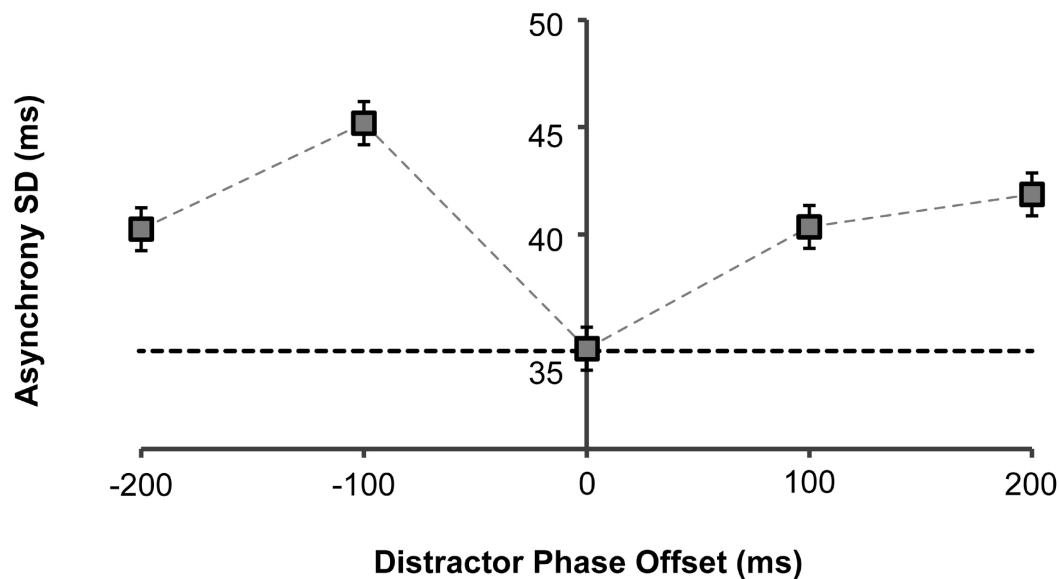
Two further analyses were carried out to compare a no-distractor condition (i.e., only the target stimulus present) with the other multiple dot conditions where there was no phase-offset between the target and distractor stimuli. As expected, we found no significant effect of the number of distractors on the mean asynchrony ( $p = 0.089$ ) or the standard deviation ( $p = 0.765$ ). Hence we can conclude that the number of distractors alone does not significantly affect synchronization to a target visual cue where there is no phase-offset applied.

## Discussion

In this study, we investigated how we synchronize our movements in time with a visually oscillating target cue in the presence of same-modality distractor cues. Participants were instructed to synchronize oscillatory arm movements in time with the target cue, while distractors varied in phase (either lagging or leading the

target cue) and number. We found that, as predicted, the degree of phase-offset between distractor and target stimuli significantly affected the asynchrony of the participants' movements to the target cue. However, contrary to expectations, an asymmetry in the distraction effect was observed, with only phase-leading distractors ( $-100$ ,  $-200$  ms) influencing the asynchrony; lagging distractors did not show any significant effect on performance. In particular, a phase offset of  $-100$  ms appeared to have a substantial impact on performance, both in terms of greater negative asynchrony and higher asynchrony variability. We further found the distraction only occurred with larger numbers of distractor stimuli surrounding the target; we saw no effect when there were only two distractor stimuli present.

The effect of distractor cues on sensorimotor synchronization performance has been investigated for combinations of auditory-auditory (Repp, 2003), auditory-visual (Repp and Penel, 2004; Hove et al., 2012; Debats et al., 2013) and auditory-proprioceptive cues (Debats et al., 2013). An asymmetry in the strength of the distraction has been observed in auditory-auditory



**FIGURE 3 | Asynchrony standard deviation (SD) as a function of distractor phase-offset to target stimulus.** There was no significant effect of number of distractors on the asynchrony SD, so results are

collapsed across this condition. Error bars represent standard errors. Dashed horizontal black line indicates baseline mean asynchrony in the no-distractor condition.

and auditory-visual conditions (Repp, 2003; Repp and Penel, 2004). In both cases, a strong influence of the auditory distractors on the asynchronies was observed when the distractors occurred earlier than the target cue, but not later. With discrete cues, this is expected: the participant's attention is captured by the early distraction events and hence draws the motor responses away from the target cue. Later distraction cues are less likely to capture attention as they occur after the motor action has been planned and executed (Repp, 2003). With a continuously present visual cue and continuous motor action however, we expected there to be no difference between a distractor being late or early in phase. We considered that the continuous signal would be a constant distraction and hence would show a symmetrical effect on the asynchronies regardless of them leading or lagging the target. The fact that we saw an asymmetry indicates that participants were still utilizing salient points in the sensory stimuli and aligning them to similarly salient anchor points in their own movement trajectories. For the visual cues, the salient points could have been, for example, the change in direction of the moving dot at the top or bottom of the sinusoidal trajectory. Indeed, it makes sense to have discrete points of reference for synchronization. On the one hand it has been shown that synchronization to continuous temporal-spatial visual cues is much easier and results in enhanced synchrony performance (Varlet et al., 2012; Armstrong et al., 2013; Armstrong and Issartel, 2014) compared to the task of timing movements to discrete visual cues (Repp and Penel, 2004; Elliott et al., 2010; Wright and Elliott, 2014). However, while the dynamic spatial element of visual information is clearly important for anticipatory timing, it would be inefficient to continuously align and correct movements at arbitrary points in the cue trajectory, just because there is the sensory information available. Evidence from this experiment and other studies (Luck and Sloboda, 2008;

Hajnal et al., 2009; Su, 2014b; Varlet et al., 2014) suggests that if we're timing movements to an external cue, we pick out discrete salient points for temporal alignment that allows us to efficiently correct movements through each repetition of the cycle. These do not have to be explicit observable events within the trajectory but can be related to derivatives of the movement such as velocity (Su, 2014b; Varlet et al., 2014) or peak acceleration (Luck and Sloboda, 2008). Similar strategies arise in the movements themselves. Producing smooth continuous movements results in the timing emerging from the movement itself [referred to as emergent, or implicit timing (Spencer et al., 2003)]. This smooth movement reduces the ability to make accurate corrections necessary for maintaining synchrony (Elliott et al., 2009a). Hence it is beneficial to timing performance to have relatively discrete (identified by a high level of jerk) features in the movement that allows event based or explicit timing (Elliott et al., 2009a). Again, in this case it is likely that proprioceptive feedback of the change of direction at the lowest point of the movement was sufficient to allow participants to synchronize their actions. These strategies of extracting discrete timing events from continuous cues and movements explains why we see a similar asymmetrical distraction effect in this task as in previous experiments that used discrete cues (e.g., an auditory metronome) and movements (finger tapping).

To understand the effect of the distractors further, we must consider the underlying attentional processes. Moving visual stimuli in the periphery attracts attention far better than static stimuli (Bartram et al., 2003). In addition, jerky motion captures attention more than smooth motion (Sunny and von Mühlenen, 2011). Our study shows that even if visual stimuli are not being attended to, the salient features of the distractor cue trajectory attract coordinated movements away from a target stimulus. It appears however, that the number of distractors



and possibly the spatial location is also important. We only observed the strong distraction effect when there were 8 or 14 distractors present. This is likely to be due to the increased salience of the distractor cues, with the large number of stimuli moving at the same phase making them increasingly difficult to ignore (Bartram et al., 2003). Equally, the salience could have been increased by the larger number of distractors completely surrounding the target dot, rather just on either side, as in the two-distractor condition. Our results therefore suggest a bottom-up stimulus driven attentional process is in place (Theeuwes et al., 2000) where the saliency of the distractor relative to the target is what draws the attention of the individual. The temporal distance of the distractors from the target is a further important factor in the strength of the distraction. With the peak distraction effect occurring when the distractors are –100 ms earlier than the target cue, it is likely a temporal window of attention (Naccache et al., 2002) around the salient event in the target cue is present. If the distractor cue event falls into this window, then it maximizes attentional capture from the target (due to the multiple distractors providing a stronger stimulus than the target). This is somewhat different to the well-documented “window of integration.” Sensory integration of temporal cues occurs when two stimuli are deemed relevant to one-another and occur close together in time (Elliott et al., 2014). In this scenario, the stimuli are integrated in a fashion that can be described under a Bayesian framework, such that the resulting combined cue becomes more reliable than either of the individual stimuli (Ernst and Banks, 2002). In a synchronization task this results in a reduced variability of the timed movements (Elliott et al., 2010). Here, we explicitly inform participants to ignore the distractor stimuli, so they are aware they are not relevant to the target.

Subsequently, we observe a high level of variability at the –100 ms offset, which is likely due to be a result of the conflict between the top-down goal of synchronizing with the target cue and the bottom-up stimulus driven effect of being attracted to the more salient distractor stimuli.

Finally, we consider these results in the context of interpersonal synchrony. Spontaneous synchrony can emerge between two individuals, often due to the strong visual cues from the partner (Richardson et al., 2007; Zivotofsky et al., 2012). Considering larger groups (e.g., crowds jumping up and down in a sports stadium), there is potentially a contextual effect on how synchrony may emerge within a group. On the one hand, an individual may be focussed on timing their movements with a known partner, in which case the movements of the remaining crowd act as distractors and hence, based on our results, are likely to weaken the coupling between the dyad. Alternatively, an individual may be moving as part of the larger crowd, in which case it would be advantageous to combine the cues from all surrounding individuals. Through sensory integration, this latter scenario should result in greater stability of synchrony within the group. In reality, a combination of these processes are likely to be present, such that within a crowd we observe an overall weak coupling across all individuals, but with strong synchrony couplings between small numbers of individuals within the crowd.

## Acknowledgments

We thank Dagmar Fraser for his assistance in the coding of the visual stimuli and Sonam Malhi for assistance with data collection. This research was funded by the Engineering and Physical Sciences Research Council [EP/I031030/1].

## References

- Armstrong, A., and Issartel, J. (2014). Sensorimotor synchronization with audio-visual stimuli: limited multisensory integration. *Exp. Brain Res.* 232, 3453–3463. doi: 10.1007/s00221-014-4031-9
- Armstrong, A., Issartel, J., Varlet, M., and Marin, L. (2013). The supplementation of spatial information improves coordination. *Neurosci. Lett.* 548, 212–216. doi: 10.1016/j.neulet.2013.05.013
- Aschersleben, G., and Prinz, W. (1995). Synchronizing actions with events: the role of sensory information. *Percept. Psychophys.* 57, 305–317. doi: 10.3758/BF03213056
- Bartram, L., Ware, C., and Calvert, T. (2003). Moticons: detection, distraction and task. *Int. J. Hum. Comput. Stud.* 58, 515–545. doi: 10.1016/S1071-5819(03)00021-1
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spat. Vis.* 10, 433–436. doi: 10.1163/156856897X00357
- Debats, N. B., Ridderikhoff, A., de Boer, B. J., and Peper, C. L. (2013). Biases in rhythmic sensorimotor coordination: effects of modality and intentionality. *Behav. Brain Res.* 250, 334–342. doi: 10.1016/j.bbr.2013.05.005
- Elliott, M. T., Welchman, A. E., and Wing, A. M. (2009a). Being discrete helps keep to the beat. *Exp. Brain Res.* 192, 731–737. doi: 10.1007/s00221-008-1646-8
- Elliott, M. T., Welchman, A. E., and Wing, A. M. (2009b). MatTAP: a MATLAB toolbox for the control and analysis of movement synchronisation experiments. *J. Neurosci. Methods* 177, 250–257. doi: 10.1016/j.jneumeth.2008.10.002
- Elliott, M. T., Wing, A. M., and Welchman, A. E. (2010). Multisensory cues improve sensorimotor synchronisation. *Eur. J. Neurosci.* 31, 1828–1835. doi: 10.1111/j.1460-9568.2010.07205.x
- Elliott, M. T., Wing, A. M., and Welchman, A. E. (2014). Moving in time: Bayesian causal inference explains movement coordination to auditory beats. *Proc. R. Soc. B Biol. Sci.* 281, 20140751. doi: 10.1098/rspb.2014.0751
- Ernst, M. O., and Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415, 429–433. doi: 10.1038/415429a
- Hajnal, A., Richardson, M. J., Harrison, S. J., and Schmidt, R. C. (2009). Location but not amount of stimulus occlusion influences the stability of visuo-motor coordination. *Exp. Brain Res.* 199, 89–93. doi: 10.1007/s00221-009-1958-3
- Helmuth, L. L., and Ivry, R. B. (1996). When two hands are better than one: reduced timing variability during bimanual movements. *J. Exp. Psychol. Hum. Percept. Perform.* 22, 278–293. doi: 10.1037/0096-1523.22.2.278
- Hove, M. J., Iversen, J. R., Zhang, A., and Repp, B. H. (2012). Synchronization with competing visual and auditory rhythms: bouncing ball meets metronome. *Psychol. Res.* 77, 388–398. doi: 10.1007/s00426-012-0441-0
- Kurgansky, A. V. (2008). Visuomotor synchronization: analysis of the initiation and stable synchronization phases. *Hum. Physiol.* 34, 289–298. doi: 10.1134/S0362119708030043
- Luck, G., and Sloboda, J. (2008). Exploring the spatio-temporal properties of simple conducting gestures using a synchronization task. *Music Percept.* 25, 225–239. doi: 10.1525/mp.2008.25.3.225
- Miura, A., Kudo, K., Ohtsuki, T., and Kanehisa, H. (2011). Coordination modes in sensorimotor synchronization of whole-body movement: a study of street dancers and non-dancers. *Hum. Mov. Sci.* 30, 1260–1271. doi: 10.1016/j.humov.2010.08.006
- Naccache, L., Blandin, E., and Dehaene, S. (2002). Unconscious masked priming depends on temporal attention. *Psychol. Sci.* 13, 416–424. doi: 10.1111/1467-9280.00474

- Noormohammadi, N., Brownjohn, J., Wing, A. M., Racic, V., Johannsen, L., and Elliott, M. T. (2011). "Effect of different cues on spectators' synchronisation, a vibration engineering approach," in *Proceedings of the 8th International Conference on Structural Dynamics*, Leuven.
- Repp, B. H. (2003). Phase attraction in sensorimotor synchronization with auditory sequences: effects of single and periodic distractors on synchronization accuracy. *J. Exp. Psychol. Hum. Percept. Perform.* 29, 290–309. doi: 10.1037/0096-1523.29.2.290
- Repp, B. H., and Penel, A. (2002). Auditory dominance in temporal processing: new evidence from synchronization with simultaneous visual and auditory sequences. *J. Exp. Psychol. Hum. Percept. Perform.* 28, 1085–1099. doi: 10.1037/0096-1523.28.5.1085
- Repp, B. H., and Penel, A. (2004). Rhythmic movement is attracted more strongly to auditory than to visual rhythms. *Psychol. Res.* 68, 252–270. doi: 10.1007/s00426-003-0143-8
- Repp, B. H., and Su, Y.-H. (2013). Sensorimotor synchronization: a review of recent research (2006–2012). *Psychon. Bull. Rev.* 20, 403–452. doi: 10.3758/s13423-012-0371-2
- Richardson, M. J., Marsh, K. L., Isenhower, R. W., Goodman, J. R. L., and Schmidt, R. C. (2007). Rocking together: dynamics of intentional and unintentional interpersonal coordination. *Hum. Mov. Sci.* 26, 867–891. doi: 10.1016/j.humov.2007.07.002
- Spencer, R. M. C., Zelaznik, H. N., Diedrichsen, J., and Ivry, R. B. (2003). Disrupted timing of discontinuous but not continuous movements by cerebellar lesions. *Science* 300, 1437–1439. doi: 10.1126/science.1083661
- Sunny, M. M., and von Mühlenen, A. (2011). Motion onset does not capture attention when subsequent motion is "smooth." *Psychon. Bull. Rev.* 18, 1050–1056. doi: 10.3758/s13423-011-0152-3
- Su, Y.-H. (2014a). Audiovisual beat induction in complex auditory rhythms: point-light figure movement as an effective visual beat. *Acta Psychol. (Amst.)* 151, 40–50. doi: 10.1016/j.actpsy.2014.05.016
- Su, Y.-H. (2014b). Peak velocity as a cue in audiovisual synchrony perception of rhythmic stimuli. *Cognition* 131, 330–344. doi: 10.1016/j.cognition.2014.02.004
- Su, Y.-H. (2014c). Visual enhancement of auditory beat perception across auditory interference levels. *Brain Cogn.* 90, 19–31. doi: 10.1016/j.bandc.2014.05.003
- Theeuwes, J., Atchley, P., and Kramer, A. F. (2000). "On the time course of top-down and bottom-up Control of Visual Attention," in *Control of Cognitive Processes: Attention and Performance XVIII*, eds S. Monsell and J. Driver (Cambridge, MA: MIT Press), 104–124.
- Varlet, M., Coey, C. A., Schmidt, R. C., Marin, L., Bardy, B. G., and Richardson, M. J. (2014). Influence of stimulus velocity profile on rhythmic visuomotor coordination. *J. Exp. Psychol. Hum. Percept. Perform.* 40, 1849–1860. doi: 10.1037/a0037417
- Varlet, M., Marin, L., Issartel, J., Schmidt, R. C., and Bardy, B. G. (2012). Continuity of visual and auditory rhythms influences sensorimotor coordination. *PLoS ONE* 7:e44082. doi: 10.1371/journal.pone.0044082
- Wright, R. L., and Elliott, M. T. (2014). Stepping to phase-perturbed metronome cues: multisensory advantage in movement synchrony but not correction. *Front. Hum. Neurosci.* 8:724. doi: 10.3389/fnhum.2014.00724
- Zivotofsky, A. Z., Gruendlinger, L., and Hausdorff, J. M. (2012). Modality-specific communication enabling gait synchronization during over-ground side-by-side walking. *Hum. Mov. Sci.* 31, 1268–1285. doi: 10.1016/j.humov.2012.01.003

**Conflict of Interest Statement:** 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.

Copyright © 2015 Booth and Elliott. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# The simultaneous perception of auditory–tactile stimuli in voluntary movement

Qiao Hao<sup>1\*</sup>, Taiki Ogata<sup>1,2</sup>, Ken-ichiro Ogawa<sup>1</sup>, Jinhwan Kwon<sup>1</sup> and Yoshihiro Miyake<sup>1</sup>

<sup>1</sup> Department of Computational Intelligence and Systems Science, Tokyo Institute of Technology, Yokohama, Japan,

<sup>2</sup> Research into Artifacts, Center for Engineering (RACE), The University of Tokyo, Kashiwa, Japan

## OPEN ACCESS

### Edited by:

Yan Bao,  
Peking University, China

### Reviewed by:

Ernst Pöppel,  
Ludwig Maximilian University  
of Munich, Germany  
Bin Zhou,  
Institute of Psychology – Chinese  
Academy of Sciences, China

### \*Correspondence:

Qiao Hao,  
Department of Computational  
Intelligence and Systems Science,  
Tokyo Institute of Technology, Midori,  
Yokohama 226-8503, Japan  
hao@myk.dis.titech.ac.jp

### Specialty section:

This article was submitted to  
Perception Science,  
a section of the journal  
Frontiers in Psychology

**Received:** 21 May 2015

**Accepted:** 07 September 2015

**Published:** 24 September 2015

### Citation:

Hao Q, Ogata T, Ogawa K, Kwon J  
and Miyake Y (2015)  
The simultaneous perception  
of auditory–tactile stimuli in voluntary  
movement. *Front. Psychol.* 6:1429.  
doi: 10.3389/fpsyg.2015.01429

The simultaneous perception of multimodal information in the environment during voluntary movement is very important for effective reactions to the environment. Previous studies have found that voluntary movement affects the simultaneous perception of auditory and tactile stimuli. However, the results of these experiments are not completely consistent, and the differences may be attributable to methodological differences in the previous studies. In this study, we investigated the effect of voluntary movement on the simultaneous perception of auditory and tactile stimuli using a temporal order judgment task with voluntary movement, involuntary movement, and no movement. To eliminate the potential effect of stimulus predictability and the effect of spatial information associated with large-scale movement in the previous studies, we randomized the interval between the start of movement and the first stimulus, and used small-scale movement. As a result, the point of subjective simultaneity (PSS) during voluntary movement shifted from the tactile stimulus being first during involuntary movement or no movement to the auditory stimulus being first. The just noticeable difference (JND), an indicator of temporal resolution, did not differ across the three conditions. These results indicate that voluntary movement itself affects the PSS in auditory–tactile simultaneous perception, but it does not influence the JND. In the discussion of these results, we suggest that simultaneous perception may be affected by the efference copy.

**Keywords:** voluntary movement, temporal simultaneity, auditory–tactile stimuli, temporal order judgment, efference copy

## Introduction

When people type quickly on a computer keyboard they usually integrate visual, auditory, and tactile information to ensure successful performance. For efficient interactions with the environment or other people, the simultaneous perception of multimodal information is important during voluntary movement, and determines the timing of multimodal events. Many previous studies have focused on the simultaneous perception of multimodal information under static experimental conditions during which participants remain immobile. However, how the timing of multimodal events is determined during voluntary movements remains largely a mystery. Although voluntary movement has been found to compress or dilate subjective time under certain

**Abbreviations:** JND, just noticeable difference; PSS, point of subjective simultaneity; SOAs, stimulus onset asynchronies; TOJ, temporal order judgment.

circumstances (Yarrow et al., 2001; Morrone et al., 2005), current knowledge about the effect of voluntary movement on auditory–tactile simultaneous perception is still unsettled. In particular, it is unclear whether voluntary movement or proprioceptive information following a movement affects the simultaneous perception of auditory and tactile stimuli.

To investigate the fundamental characteristics of simultaneous perception, simultaneity judgment (SJ) tasks (Schneider and Bavelier, 2003; Zampini et al., 2005a) or TOJ tasks (Mitrani et al., 1986; Spence et al., 2001; Zampini et al., 2003; Miller and Schwarz, 2006; Cardoso-Leite et al., 2007; Boenke et al., 2009; Van Eijk et al., 2009; Kwon et al., 2014) are often used. In a SJ task, two stimuli are presented at various SOAs and the participants are asked to indicate whether the two stimuli are simultaneous or not. In a TOJ task, the participants are required to judge the temporal order of the two stimuli. These tasks have revealed that people tend to perceive different modal stimuli as occurring simultaneously when they are presented with a short lag (Slutsky and Recanzone, 2001; Lewald and Guski, 2003; Kayser et al., 2008; Shi et al., 2008; Nishi et al., 2014). More specifically, the PSS differs from the point of physical simultaneity. Furthermore, temporal resolution is usually evaluated by JND, which represents difference threshold of SJ or TOJ task, with a lower JND indicating higher temporal resolution, and vice versa. JNDs differ for different combinations of multimodal information types (Keetels and Vroomen, 2005, 2008; Zampini et al., 2005b).

Some previous studies have shown that voluntary movements affect the PSSs and/or JNDs between visual–tactile stimuli (Vogels, 2004; Shi et al., 2008) and between auditory–tactile stimuli (Kitagawa et al., 2009; Frissen et al., 2012; Nishi et al., 2014) in SJ and TOJ tasks compared with conditions without voluntary movement. To investigate the effect of voluntary movement on simultaneous perception, the effect of proprioceptive sensation attending the movement must be separated from that of voluntary movement. If PSS and/or JND changes are observed even when the proprioceptive information effect is excluded, we can say that the voluntary movement itself has some influence on simultaneous perception. Therefore, voluntary, involuntary, and no movement conditions were used in previous studies (Kitagawa et al., 2009; Frissen et al., 2012; Nishi et al., 2014). Because a device moved the participants' body parts in the involuntary movement condition in the previous studies, the involuntary movement was attended by proprioceptive information. Therefore, the comparison between the involuntary and no movement conditions showed the effect of the proprioceptive information, and the comparison between the voluntary and involuntary movement conditions revealed the effect of voluntary movement exclusive of proprioceptive information.

However, those investigations of the effect of voluntary movements on the PSSs and JNDs in the auditory–tactile TOJ tasks reported contradictory results (Table 1, Effect on PSS and Effect on JND rows). Kitagawa et al. (2009) found that voluntary movement did not affect the PSS, whereas Nishi et al. (2014) found that voluntary movement caused the PSS to be associated with a preceding auditory stimulus. In addition, Frissen et al.

**TABLE 1 | Comparison of methods and results among three previous studies of the effect of voluntary movement on auditory–tactile TOJ tasks.**

Study	Kitagawa et al., 2009	Frissen et al., 2012	Nishi et al., 2014
Conditions of movement	Vol Inv Pr No	Vol Inv No	Vol Inv No
Predictability of the stimulus	Predictable	Not predictable	Predictable
Moving body part	Right index finger	Forearm, hand, and finger	Right index finger
Effect on PSS	N.S.	A shifted to T (Inv. movement)	T shifted to A (Vol. movement)
Effect on JND	L (Vol. movement)	N.S.	L (Vol. movement)

"Vol," "Inv," "Pr," and "No" indicate the voluntary, involuntary, predictable, and no movement conditions. For the effect on PSS, "N.S." means no significant difference. "A shifted to T" means the PSS in the involuntary movement condition shifted from the auditory stimulus first as in the no movement condition to the tactile stimulus first, where "A" and "T" indicate the auditory and tactile stimuli. "T shifted to A" means the PSS in the voluntary movement condition shifted from the tactile stimulus first as in the no movement condition to the auditory stimulus first. For the Effect on JND, "L" and "H" indicate that the JND was improved (lower JND) or impaired (higher JND) by voluntary movement.

(2012) found that involuntary movement caused the PSS to be associated with a preceding tactile stimulus. On the other hand, although Frissen et al. (2012) observed no effect of voluntary movement on the JND, Kitagawa et al. (2009) and Nishi et al. (2014) reported that voluntary movement improved the JND. These differing results may have been caused by unexpected effects associated with the different experimental methods used in the previous studies, such as the predictability of the stimulus and the amount of movement (Table 1, Predictability of the stimulus and Moving body part rows). For instance, a predictable stimulus could directly improve the JND (Petrini et al., 2009; Yokoyama et al., 2009; Vroomen and Stekelenburg, 2010). The spatial information in large-scale movement could obscure the effect of involuntary movement on the PSS.

Kitagawa et al. (2009) conducted the TOJ task under four conditions: voluntary, involuntary, predictable, and no movement (Table 1, Conditions of movement row). The participants pressed the button voluntarily and involuntarily with their fingers in the voluntary and involuntary movement conditions, respectively. The predictable condition was designed to enable participants to predict the occurrence of the stimulus in the TOJ task. The authors concluded that voluntary movement improved the participants' JND, because there was no improvement in the JNDs of the involuntary, predictable, and no movement conditions. However, in Kitagawa et al.'s (2009) procedure, tactile stimulation was generated as a result of voluntary finger movement. This effect induced the participants to predict the onset of the tactile stimulus (Table 1, Predictability of the stimulus row), and improved the JND of the voluntary movement condition (Table 1, Effect on JND row). Nishi et al. (2014) conducted the TOJ task under three conditions: voluntary, involuntary, and no movement (Table 1, Conditions of movement row). The authors used a device that presented tactile stimulus during voluntary finger movement to solve the problem in Kitagawa et al.'s (2009) procedure. Nevertheless, this



prediction effect on the improvement in the JND associated with voluntary movement also occurred in Nishi et al.'s (2014) study (Table 1, Effect on JND and Predictability of the stimulus rows), because the tactile stimulus was always presented 500 ms after the finger movement. It was easier to predict the stimulus onset in the voluntary movement condition.

This predictability of stimulus onset did not appear in the Frissen et al.'s (2012) study. The authors used a device that presented the tactile stimulus for the TOJ task at random interval in the voluntary movement condition, to prevent the stimulus predictability (Table 1, Predictability of the stimulus row). As a result, Frissen et al. (2012) reported that voluntary movement did not affect the JND (Table 1, Effect on JND row). This result suggests that the predictability of the stimulus improved the JNDs both in the Kitagawa et al.'s (2009) and Nishi et al.'s (2014) studies. On the other hand, Frissen et al. (2012) reported that the tactile stimulus occurring first was perceived as the PSS in the involuntary movement condition (Table 1, Effect on PSS row). However, the spatial information in large-scale movement (Table 1, Moving body part row) could have obscured the effect of movement on the PSS in Frissen et al.'s (2012) study. The large-scale movement could lead to a tactile version of a flash-lag effect (FLE; Kitagawa et al., 2005). In this phenomenon, observers perceived a flash lag behind a spatially aligned moving stimulus (Nijhawan, 2002).

Therefore, the aim of the present study was to investigate whether only voluntary movement alone affects the simultaneous perception of auditory and tactile stimuli, that is, independent of the effects of stimulus predictability and the spatial information inherent in large-scale movement (which were thought to be the causes of the divergent results in previous studies). We hypothesized that the PSS would shift from the tactile stimulus first in the involuntary movement or no movement condition to the auditory stimulus first in the voluntary movement condition. Thus, we randomized the interval between the start of movement and the first stimulus to prevent the participants from predicting the stimulus onset. In addition, we used small-scale movement to minimize the effect of spatial information on perceived simultaneity.

## Materials and Methods

### Participants

Eighteen participants (three females and 15 males, mean age: 23 years, age range: 21–28 years) completed the experiment. All of the participants were right-handed, with normal auditory thresholds and senses of touch, and they did not exhibit any difficulty moving their right index fingers. Informed consent was obtained in writing from all the participants prior to their participation in the experiment. The participants were paid for their participation, and the experiment was approved by the ethics committee of the Tokyo Institute of Technology.

### Apparatus and Stimuli

The auditory stimulus was a sinusoidal wave sound (2000 Hz, 50 dB, 10 ms) presented in both ears simultaneously via

earphones (Radius HP-RHF41; Machida, Tokyo, Japan). The tactile stimulus was an impulse force (5 N, 10 ms, rectangular pulse) provided by a PHANTOM Desktop haptic device (SensAble Technologies, Woburn, MA, USA) and orthogonal to the finger movement. The 10 ms duration for auditory and tactile stimuli was selected to avoid a problem of the procedure in the Frissen et al.'s (2012) study. In that study, the duration of the auditory stimulus (100 ms) was considerably longer than that of the tactile stimulus (10 ms). Stimulus duration has been found to create an attractor effect on the PSS in audiovisual TOJ task (Boenke et al., 2009). In other words, with increasing stimulus duration, positive PSSs shift toward negative values (because the visual stimulus must precede the auditory stimulus for simultaneous perception), and negative PSSs shift toward positive values. Hence, we used the same duration for the two stimuli. The timing of the two presentations and the movement of the device were controlled to within an error margin of 1 ms. These sensory stimulation systems were operated by computer programs installed on a PC workstation (HP xw4600/CT; Hewlett-Packard, Palo Alto, CA, USA), and were developed with the Open Haptics software development toolkit (SensAble Technologies) on the Microsoft Visual C++ 2008 platform (Microsoft, Redmond, WA, USA).

### Task and Conditions

For the TOJ task, auditory–tactile stimulus pairs were presented to participants with varying SOAs (intervals between the within-pair onsets of the auditory and tactile stimuli), and the participants judged the temporal order of the two stimuli. The SOAs were  $\pm 240$ ,  $\pm 120$ ,  $\pm 60$ ,  $\pm 30$ , and 0 ms (where the positive values indicate that the auditory stimulus was presented before the tactile stimulus, and vice versa). We chose these SOAs to improve the procedures in the Frissen et al.'s (2012) study. In that study, they used a 75 ms increment between their SOAs (300, 225, 150, 75, and 0 ms), which is a little larger than the increments used in previous auditory–tactile integration studies (Zampini et al., 2005b; Fujisaki and Nishida, 2009). Thus, we used a smaller increment for our SOAs.

There were three conditions in this experiment: voluntary, involuntary, and no movement. The involuntary movement trajectory was reproduced from voluntary movement data collected in the preliminary experiments. The mean rate of movement of the participants' fingers was 81.08 mm/s ( $SD = 7.33$ ) in the voluntary movement condition and  $\sim 78.23$  mm/s ( $SD = 1.44$ ) in the involuntary movement condition (as guided by the haptic device). The participants were seated in a darkened, sound-attenuated room in front of the stimulation systems, with the palmar side of their right index fingers held on the haptic device. They also wore sound-insulating earmuffs over their earphones and an eye mask to eliminate the confounding effect of visual stimuli during the experiment (Figure 1). In each condition, the participants were asked to indicate the temporal order of the auditory and tactile stimuli by using the Z and X keys on a keyboard. The Z indicated that the auditory stimulus occurred first and the X indicated that the tactile stimulus occurred first.



FIGURE 1 | Experimental environment.

## Procedure

### Voluntary Movement Condition

For each trial (Figure 2A), the participants voluntarily and naturally began to move their right index fingers from right to left at their own pace. As they did, a cue sound (distinct from the target auditory stimulus) indicated that the TOJ task was forthcoming. The first stimulus (either tactile or auditory) was then presented with a random delay of 600–700 ms after the cue sound onset. The second stimulus (auditory or tactile, whichever was not presented first) followed the first stimulus, offset by one of the nine SOAs previously mentioned. The participants then indicated which stimulus occurred first using a two-alternative

forced-choice test (as described above). If the participants did not move at a speed of 50–110 mm/s, they were given one more trial, randomly chosen from the remaining trials.

### Involuntary Movement Condition

Similar to the voluntary movement condition, the haptic device randomly started to move the participants' right index fingers from right to left for 500 to 1000 ms, to reproduce the variance in the onsets of voluntary movements in the preliminary experiments. The procedure for evaluating the temporal order of the two stimuli and the SOA values were the same as in the voluntary movement condition. A speed of 76 mm/s for the finger movement was set for each trial (Figure 2B), because this was considered to be a comfortable speed and representative of normal surface exploration.

### No Movement Condition

The participants in the no movement condition remained stationary throughout each trial, with the palmar side of their right index fingers held on the haptic device (Figure 2C). The first stimulus (either tactile or auditory) was presented with a random delay (600–700 ms) after the cue sound onset. The presentation of the second stimulus and the procedure for evaluating the temporal order of the two stimuli were the same as in the voluntary and involuntary movement conditions. We used the 600–700 ms interval between the cue sound onset and the first stimulus to improve the procedure used in Nishi et al.'s (2014) study. In that study, the interval between the cue sound onset and the first stimulus was 1800–3300 ms in the no movement condition, whereas it was 500 ms between the cue sound onset (or the start of movement) and the tactile stimulus for all trials in the voluntary and involuntary movement conditions. This may have affected the comparisons among the conditions, because the

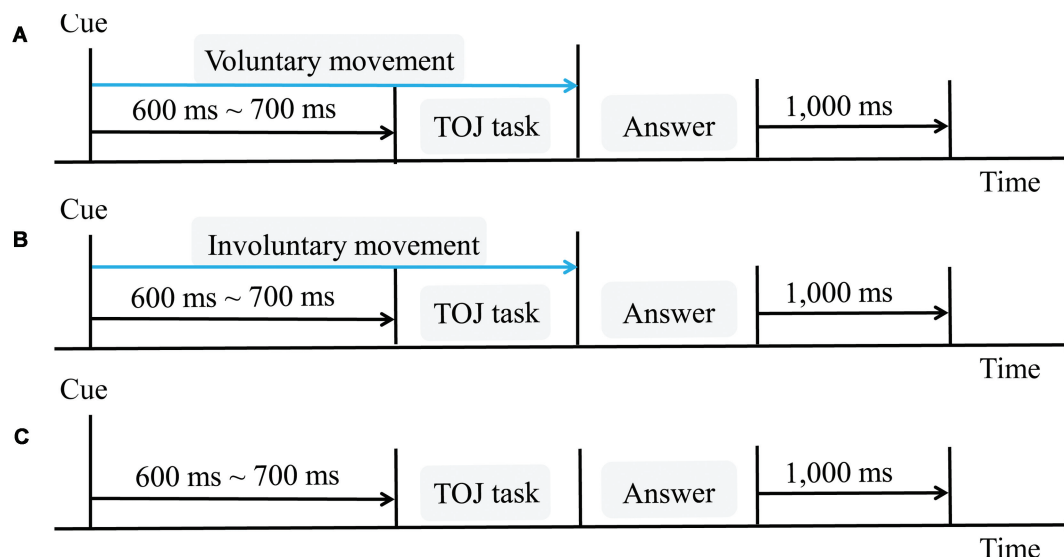


FIGURE 2 | Schematic flow chart for one trial in each of the three conditions. (A) Voluntary movement condition, in which participants voluntarily started to move their right index fingers; (B) involuntary movement condition, in which the haptic device moved the participants' right index fingers; (C) no movement condition. The interval between the cue and the TOJ task was randomly set from 600 to 700 ms. The interval between trials was 1000 ms.

different cue-target intervals activate distinct brain areas (Coull et al., 2000), affect temporal discrimination, and influence early perceptual processing (Sanders and Astheimer, 2008).

Each participant completed three blocks of trials in each of the conditions in the present experiment. The conditions were presented in a random order, and the participants were blind to the order of the conditions. Each block consisted of 45 trials, comprising five trials for each SOA, randomly selected from the following values:  $\pm 240$ ,  $\pm 120$ ,  $\pm 60$ ,  $\pm 30$ , and 0 ms. Thus, each participant completed 405 trials. The interval between trials was 1000 ms in each condition, and white noise was played in the background to effectively mask any sounds made by the haptic device. It took  $\sim 5$  min for the participants to complete one block of trials. They were given several minutes of rest between blocks, according to their preferences. The order of the conditions was counterbalanced, and the entire procedure took  $\sim 2$  h. To accustom the participants in the voluntary movement condition to the appropriate finger speeds, they each completed a practice run of ten trials in which only the tactile stimulus was presented. To eliminate this compound effect (e.g., sensitization of the tactile channel), the participants were given 2–3 min of rest before each block of trials in the voluntary movement condition. Additionally, the participants were asked to pay constant attention to the tactile stimulus to control for the prior entry effect (Shore et al., 2001; Spence et al., 2001; Kitagawa et al., 2005; Zampini et al., 2005c), which facilitates the processing of an attended stimulus relative to an unattended stimulus.

For each trial in the practice sessions, the participants were asked to close their eyes and judge the order of the two stimuli and then open their eyes to see the feedback on the computer screen. With no information about the forthcoming condition, they completed 45, 20, and 20 trials in the voluntary, involuntary, and no movement conditions, respectively. The orders of the trials were counterbalanced, and the SOA was randomly chosen from  $\pm 240$ ,  $\pm 120$ , and  $\pm 60$  ms. In addition, the short interval (600–700 ms) between the onset of the movement and the TOJ task may have produced a strong interaction between the tactile signals elicited by the onset of the movement and by the tactile stimulus in the TOJ task. Thus, there appears to be a risk that the results of this study may be unclear. In fact, movement onset has been found to impair the temporal order threshold immediately following operant actions, but then reverts in the later action-effect interval (450–850 ms; Wenke and Haggard, 2009). Furthermore, the potential strong interaction did not appear to affect the tactile TOJ tasks in studies by Hermosillo et al. (2011) or Nishikawa et al. (2015), in which they used short intervals between the onset of movements and TOJ tasks. Therefore, the possibility of a strong interaction does not threaten the results of this study.

## Data Analysis

We used MATLAB Statistics Toolbox (MathWorks, Natick, MA, USA) for the statistical regression calculations and graphic representation of the results. First, we calculated for each SOA the proportion of the answers, in which the auditory stimulus was perceived first. Then, logistic regressions were conducted using a generalized linear model with the ratio data for each

condition. Psychometric curves were fitted to the distribution of the mean TOJ data for the voluntary, involuntary, and no movement conditions, as shown in **Figure 3**.

The values of the PSS and JND were calculated for each participant in the regression analysis based on three equations (Finney, 1952):

$$y = \frac{1}{1 + e^{-\frac{\alpha - x}{\beta}}} \quad (1)$$

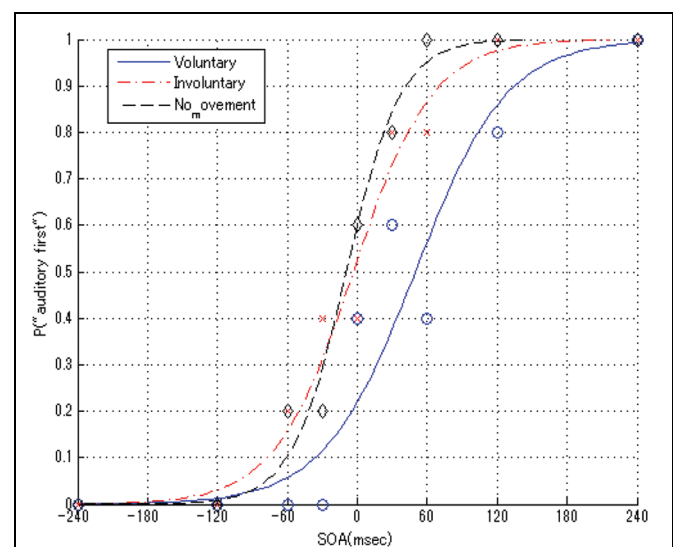
$$\text{PSS} = \alpha \quad (2)$$

$$\text{JND} = \frac{x_{75} - x_{25}}{2} = \beta \log 3 \quad (3)$$

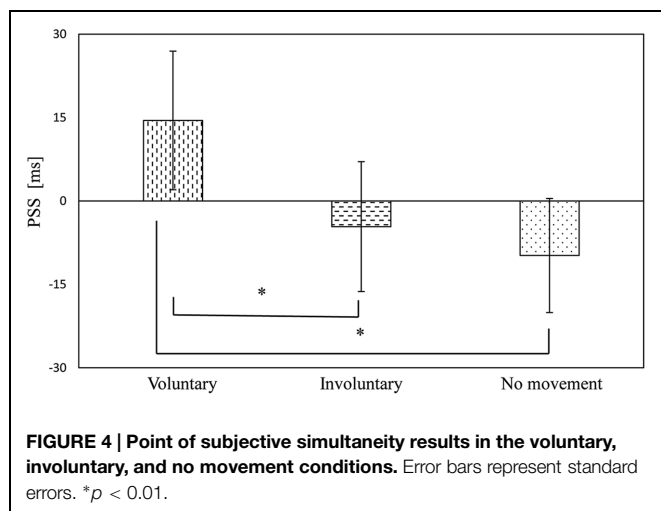
Here,  $\alpha$  represents the estimated PSS,  $x$  denotes the SOA,  $\beta$  is related to the JND, and  $x_p$  represents the SOA with  $p$  as the percent of “auditory first” responses. Then, a statistical analysis of the data was conducted to obtain the mean and standard error values for each condition.

## Results

The PSSs of the voluntary, involuntary, and no movement conditions were 14.5 ms ( $SE = 12.5$ ),  $-4.6$  ms ( $SE = 11.7$ ), and  $-9.8$  ms ( $SE = 10.3$ ), respectively, as shown in **Figure 4**. A one-way repeated measures analysis of variance (ANOVA) with movement condition as a factor showed a significant effect [ $F(2,34) = 12.74$ ,  $p < 0.001$ ]. Subsequently, Bonferroni-Holm paired  $t$ -tests revealed significant differences between the voluntary and involuntary movement conditions ( $p = 0.001$ ), and between the voluntary and no movement conditions ( $p = 0.008$ ). There was no significant difference between the involuntary and



**FIGURE 3 | Average psychometric functions among all blocks in the voluntary, involuntary, and no movement conditions for one participant.** Positive SOA values mean that the auditory stimulus was presented before the tactile stimulus, and vice versa.

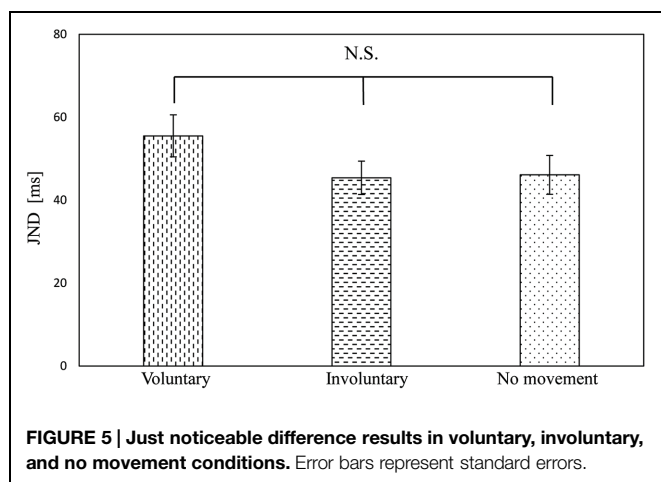


no movement conditions ( $p = 0.70$ ), as shown in **Figure 4**. The magnitude of the effect size in the PSS ( $\eta^2 = 0.43$ ) was large (Cohen, 1988).

The JNDs of the voluntary, involuntary, and no movement conditions were 55.5 ms ( $SE = 5.1$ ), 45.4 ms ( $SE = 4.0$ ), and 46.1 ms ( $SE = 4.7$ ), respectively. A one-way repeated measures ANOVA with movement condition as a factor was not significant [ $F(2,34) = 2.28$ ,  $p = 0.12$ ], with  $p = 0.26$  between the voluntary and involuntary movement conditions,  $p = 0.30$  between the voluntary and no movement conditions, and  $p = 1.0$  between the involuntary and no movement conditions, as shown in **Figure 5**. The magnitude of the effect size for the JND ( $\eta^2 = 0.12$ ) was medium (Cohen, 1988).

## Discussion

The aim of this study was to isolate the potential impacts of methodological differences on the results of previous studies and investigate the effect of voluntary movement on the simultaneous perception of auditory and tactile stimuli in a TOJ task. In



the present study, the potential effect of predictability on JNDs in Kitagawa et al.'s (2009) and Nishi et al.'s (2014) studies was removed by randomizing the interval between the start of movement and the first stimulus in the voluntary movement condition. Furthermore, we minimized the potential effect of the spatial information associated with large-scale movement on the PSS of involuntary movement condition (which was a problem in the Frissen et al.'s (2012) study) by using small-scale movement.

The results of this study replicated the effect of voluntary movement on the PSS (Nishi et al., 2014) and the JND (Frissen et al., 2012) in previous studies. In this study, we found that there was a significant shift in the PSS of the voluntary movement condition relative to the PSS of the involuntary and no movement conditions. There was no significant difference in the PSS between the involuntary and no movement conditions. The JND was not influenced by voluntary movement compared with the other two conditions. We discuss these differences in more detail below.

## Effect of Voluntary Movement on PSS

**Table 2** shows the PSS results of the previous and present studies. The PSS shift associated with involuntary movement in the Frissen et al.'s (2012) and Nishi et al.'s (2014) studies was not observed in the present study (Inv–No column). This result suggests that in the Frissen et al.'s (2012) study, the spatial information of the large-scale movement significantly caused the PSS shift in the involuntary movement condition because the present study minimized the effect of spatial information in the involuntary movement condition. In addition, the lack of short-range SOAs in Frissen et al.'s (2012) study may conceal the difference between voluntary and no movement conditions (see Materials and Methods; Vol–No column). The different stimulus durations would also partially confound the interpretation of the PSSs in their results (see Materials and Methods). This result also suggests that in the Nishi et al.'s (2014) study, the PSS shift associated with involuntary movement was caused by the different intervals, which were between the start of movement and the tactile stimulus in the involuntary movement condition, and between the cue sound onset and the first stimulus in the no movement condition, respectively (see Materials and Methods). This effect caused by different intervals did not occurred in the present study, because we used the same interval between cue sound onset and the first stimulus throughout in the three conditions. The reasoning for this is as follows. First, the long cue-target intervals activate the areas of the brain involved in motor preparation, which are distinct from those activated by short cue-target intervals (Coull et al., 2000). Second, it has been found that temporal discrimination is better between 500 and 1000 ms than it is between 1000 and 1500 ms, with sounds beginning 500, 1000, and 1500 ms after the onset of a fixation point (as at the start of a trial; Sanders and Astheimer, 2008). Sanders and Astheimer (2008) found that this flexibility of temporally selective attention affects early perceptual processing.

**Table 2** also shows that voluntary movement shifts the PSS of an auditory–tactile TOJ from the tactile stimulus being first to



**TABLE 2 | Comparison among PSS results.**

Condition	PSS			Difference		
	Vol	Inv	No	Vol-Inv	Vol-No	Inv-No
Frissen et al., 2012	12	-45	4	*	N.S.	*
Nishi et al., 2014	5.3	-13.1	-37.7	*	**	**
This study	14.5	-4.6	-9.8	*	*	N.S.

"Vol," "Inv," and "No" indicate voluntary, involuntary, and no movement conditions, respectively. "Vol-Inv," "Vol-No," and "Inv-No" indicate the differences between the respective conditions. A negative PSS represents the presentation of the tactile stimulus before the auditory stimulus. \* $p < 0.01$ ; \*\* $p < 0.001$ .

the auditory stimulus being first (Vol-Inv and Vol-No columns), but that a proprioceptive sensation does not affect the PSS (Inv-No column). One possible explanation for the accelerated processing speed of the tactile stimulus by voluntary movement is efference copy. Efference copy, which is a copy of the motor command, is generated in the presupplementary motor cortex and the premotor cortex (Tanji and Mushiake, 1996). Evidence from three lines of research—functional magnetic resonance imaging (fMRI) experiments in humans (Cui et al., 2014), the activation of Brodmann area 2 (BA2) neurons in activity preceding the active movements of monkeys (Weber et al., 2011), and neurons recorded in the somatosensory cortex (SI, BA2 in particular) that only discharge during voluntary movements (London and Miller, 2013)—indicates that the efference copy can significantly influence the primary somatosensory cortices. The somatosensory cortex, which is also modulated by the premotor cortex during voluntary movements without proprioceptive feedback (Christensen et al., 2007), is an area of the brain that processes input from the various systems of the body, and is sensitive to touch. In addition, the efference copy is sent to the posterior parietal cortex (Desmurget et al., 2009), where tactile events are localized in external space (Azañon et al., 2010). Therefore, the efference copy of a voluntary movement may affect the processing speed of the tactile stimulus in the TOJ task used in this study.

A second possible explanation for the accelerated processing of the tactile stimulus in the voluntary movement condition is that the participants experienced the illusion of a self-generated tactile stimulus (as a kind of causal belief), which only occurs with self-paced voluntary movements. Based on this action-effect prediction (Waszak et al., 2012), the efference copies of the self-generated tactile stimulus and voluntary movement affected the processing speed of the tactile stimulus and then changed the PSS. This effect is identical to that in the Directions into Velocities of Articulators (DIVA) model in the online control of speech production. In the DIVA model, an efference copy of the motor command was found to be useful for motor preparation, and the auditory efference copy predicted the possible auditory outcome (Guenther et al., 2006). In addition, there is neurophysiological evidence of the human brain deploying efference copies in the somatosensory and auditory cortices in finger tapping and speech production tasks, respectively (Rauschecker and Scott, 2009; Tian and Poeppel, 2010).

In addition, voluntary movement may not only affect the processing speed of the tactile stimulus but also influence the TOJ task itself. Neural imaging evidence from fMRI studies has identified the activation of the temporal parietal junction (TPJ). This evidence was reported for TOJ tasks between two visual stimuli (Davis et al., 2009) and between two tactile stimuli (Takahashi et al., 2013), as well as between auditory and visual stimuli (Adhikari et al., 2013). The efference copy of a motor command is sent to the posterior parietal cortex (Desmurget et al., 2009), and the close relationship between the locations of the posterior parietal cortex and the TPJ proposed by Nishi et al. (2014) led us to infer that voluntary movement could influence the TOJ task itself.

Another reason why the shift of PSS occurring in voluntary movement (Table 2, Vol-Inv and Vol-No columns) may be related to the prior entry effect (Shore et al., 2001; Spence et al., 2001; Kitagawa et al., 2005; Zampini et al., 2005c). Both endogenous and exogenous attention to stimuli may change the PSS. In the present study, endogenous and exogenous attention may have been mixed. First, voluntary movement may enhance endogenous attention to tactile stimuli. The prior entry effect may have occurred and caused the PSS shift in the voluntary movement condition. Second, voluntary movement may decrease auditory exogenous attention, assuming that the auditory cue at the start of the trial increased auditory exogenous attention. We asked the participants to pay attention to a tactile stimulus in the three conditions to control for the prior entry effect (endogenous attention to tactile stimuli). However, voluntary movement may increase endogenous attention to tactile stimuli and decrease the effect of auditory exogenous attention. This attention shift may accelerate the speed of tactile processing and/or reduce the speed of auditory processing in the voluntary movement condition, which would lead to a PSS shift.

## Effect of Voluntary Movement on JND

Table 3 shows the JND results of the previous and present studies. There were significant differences between the involuntary movement or no movement condition and the voluntary movement condition in Nishi et al.'s (2014) study, but there was no difference among the three conditions both in the present study and in Frissen et al.'s (2012) study. That is, both this study and Frissen et al.'s (2012) study failed to find an effect of voluntary movement on the JND.

**TABLE 3 | Comparison among JND results.**

Condition	JND			Difference		
	Vol	Inv	No	Vol-Inv	Vol-No	Inv-No
Frissen et al., 2012	114	94	102	N.S.	N.S.	N.S.
Nishi et al., 2014	46.8	59.2	66.1	*	**	N.S.
This study	55.5	45.4	46.1	N.S.	N.S.	N.S.

"Vol," "Inv," and "No" indicate the voluntary, involuntary, and no movement conditions. "Vol-Inv," "Vol-No," and "Inv-No" indicate the differences between the respective conditions. \* $p < 0.01$ ; \*\* $p < 0.001$ .

The present results suggest that the improved JND in Nishi et al.'s (2014) study were caused by the predictability of the stimulus. In their experiments, the tactile stimulus was always presented 500 ms after the finger movement in the voluntary movement condition. This could have allowed the participants to predict the occurrence of the stimulus and improve their JNDs (Petrini et al., 2009; Yokoyama et al., 2009; Vroomen and Stekelenburg, 2010). This stimulus predictability occurs only in the voluntary movement condition, because the JND in the involuntary movement condition, in which tactile stimulus was always presented 500 ms after the finger movement, did not differ from the JND in the no movement condition. On the other hand, the JND values in the present study are lower than those reported by Frissen et al. (2012). This means that the temporal window for auditory–tactile integration was narrower in this study than in the Frissen et al.'s (2012) study. We included a practice session in our experiment before the formal experimental trials to familiarize the participants with the TOJ task. Furthermore, the participants had additional practice in the voluntary movement condition to ensure appropriate finger speeds. Therefore, relative to the participants in Frissen et al.'s (2012) study, our participants were well-trained prior to the experimental conditions. The difference in JND values between the present study and Frissen et al.'s (2012) study was consistent with the findings of Hirsh and Sherrick (1961), in which well-trained participants performed better than less well-trained ones.

## Limitations

This study has some limitations. First, the practice session before the voluntary movement condition in which only tactile stimuli were presented may have an effect on the results (i.e., sensitization of the tactile channel in the voluntary movement condition). To eliminate the confounding effect of this practice session, the participants were given 2–3 min rest before each block of trials in the voluntary movement condition. We believe that this eliminated the effect of the voluntary movement condition practice runs on the observed results of the JND and PSS. First, according to a previous study (Hirsh and Sherrick, 1961), the more that people practice, the more their JNDs improve. However, JND did not improve in the voluntary movement condition in this study. This suggests that the potentially confounding effect of practice was well-controlled in this study. Second, according to another previous study (Zampini et al., 2005b), the amount of practice does not affect the PSS in auditory–tactile stimuli TOJ task. Therefore, there is no reason to believe that the practice session prior to the voluntary movement condition impacted the PSS result. However, further investigation may be necessary on this issue.

There may be a second limitation of this study related to stimulus intensity. Boenke et al. (2009) showed that stimulus intensity plays a role in the temporal perception of auditory–visual stimulus pairs. We used a stronger tactile stimulus in the present study than Frissen et al.'s (2012) study, and thus the strength of the tactile stimulus may have interacted strongly with

the voluntary movement in our experiment. In future work, it would be interesting to investigate how the relationship between the strength of the tactile stimulus and voluntary movement affects simultaneous perception.

Finally, the ratio of male to female participants in this study was 5:1, which may limit the generalizability of the results. Although previous research has shown that there are no gender effect on two tactile TOJ task in the uncrossed arms condition (Cadieux et al., 2010) or on the temporal order threshold of two types of paired tones stimuli (Bao et al., 2013), it is unknown whether a gender difference exists in multimodal integration. Thus, it would be useful in future research to include more female participants to determine whether there is gender difference in the multimodal integration of auditory and tactile information in TOJ task.

## Conclusion

The purpose of this study was to investigate the effect of voluntary movement on auditory–tactile simultaneous perception, controlling for the effects of stimulus predictability, spatial information associated with large-scale movement, and other methodological problems (see Materials and Methods) found in previous studies (Kitagawa et al., 2009; Frissen et al., 2012; Nishi et al., 2014). Auditory–tactile TOJ tasks were conducted in voluntary, involuntary, and no movement conditions. The PSS in the voluntary movement condition shifted from the tactile stimulus being first in the involuntary movement or no movement condition to the auditory stimulus being first. JNDs did not differ across the three conditions. These results reveal that voluntary movement changes the PSS rather than the JND, but proprioceptive information does not affect the simultaneous perception of auditory and tactile stimuli.

Up until now, many studies of the simultaneous perception of multimodal information have focused on the no movement condition, in which participants simply receive information from the environment. However, we routinely act voluntarily on the environment and receive sensory feedback from the environment, with these two events together defining the moment. Therefore, it is necessary to study the simultaneous perception of multimodal information in voluntary movements, and not just in static (no movement) situations.

## Acknowledgments

We would like to thank Mr. Leo Ota of the Tokyo Institute of Technology for his valuable discussions and Mr. Yuki Hirobe of the Tokyo Institute of Technology for his assistance with programming. This study was supported by Japan Society for the Promotion of Science (KAKENHI) on Scientific Research on Innovative Areas Program Grant Number 26560114 and Scientific Research (A) Program Grant Number 15H01771.

## References

- Adhikari, B. M., Goshorn, E. S., Lamichhane, B., and Dhamala, M. (2013). Temporal-order judgment of audiovisual events involves network activity between parietal and prefrontal cortices. *Brain Connect.* 3, 536–545. doi: 10.1089/brain.2013.0163
- Azañon, E., Longo, M. R., Soto-Faraco, S., and Haggard, P. (2010). The posterior parietal cortex remaps touch into external space. *Curr. Biol.* 20, 1304–1309. doi: 10.1016/j.cub.2010.05.063
- Bao, Y., Szysmaszek, A., Wang, X., Oron, A., Pöppel, E., and Szelag, E. (2013). Temporal order perception of auditory stimuli is selectively modified by tonal and non-tonal language environments. *Cognition* 129, 579–585. doi: 10.1016/j.cognition.2013.08.019
- Boenke, L. T., Deliano, M., and Ohl, F. W. (2009). Stimulus duration influences perceived simultaneity in audiovisual temporal order judgment. *Exp. Brain Res.* 198, 233–244. doi: 10.1007/s00221-009-1917-z
- Cadieux, M. L., Barnett-Cowan, M., and Shore, D. I. (2010). Crossing the hands is more confusing for females than males. *Exp. Brain Res.* 204, 431–446. doi: 10.1007/s00221-010-2268-5
- Cardoso-Leite, P., Gorea, A., and Mamassian, P. (2007). Temporal order judgment and simple reaction times: evidence for a common processing system. *J. Vis.* 7, 1–14. doi: 10.1167/7.6.11
- Christensen, M. S., Lundbye-Jensen, J., Greetsen, S. S., Petersen, T. H., Paulson, O. B., and Nielsen, J. B. (2007). Premotor cortex modulates somatosensory cortex during voluntary movements without proprioceptive feedback. *Nat. Neurosci.* 10, 417–419.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*, 2nd Edn. Hillsdale, NJ: Erlbaum.
- Coull, J., Frith, C., Büchel, C., and Nobre, A. (2000). Orienting attention in time: behavioral and neuroanatomical distinction between exogenous and endogenous shifts. *Neuropsychologia* 38, 808–819. doi: 10.1016/S0028-3932(99)00132-3
- Cui, F., Arnstein, D., Thomas, R. M., Maurits, N. M., Keysers, C., and Gazzola, V. (2014). Functional magnetic resonance imaging connectivity analyses reveal efference-copy to primary somatosensory area, BA2. *PLoS ONE* 9:e84367. doi: 10.1371/journal.pone.0084367
- Davis, B., Christie, J., and Rorden, C. (2009). Temporal order judgments activate temporal parietal junction. *J. Neurosci.* 29, 3182–3188. doi: 10.1523/JNEUROSCI.5793-08.2009
- Desmurget, M., Reilly, K. T., Richard, N., Szathmari, A., Mottolese, C., and Sirigu, A. (2009). Movement intention after parietal cortex stimulation in humans. *Science* 324, 811–813. doi: 10.1126/science.1169896
- Finney, D. J. (1952). *Probit Analysis: A Statistical Treatment of the Sigmoid Response Curve*. England: Cambridge University Press.
- Frissen, I., Ziat, M., Campion, G., Hayward, V., and Guastavino, C. (2012). Effects of voluntary movements on auditory–haptic and haptic–haptic temporal order judgments. *Acta Psychol. (Amst.)* 141, 140–148. doi: 10.1016/j.actpsy.2012.07.010
- Fujisaki, W., and Nishida, S. (2009). Audio-tactile superiority over visuo-tactile and audio-visual combinations in the temporal resolution of synchrony perception. *Exp. Brain Res.* 198, 245–259. doi: 10.1007/s00221-009-1870-x
- Guenther, F. H., Ghosh, S. S., and Tourville, J. A. (2006). Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain Lang.* 96, 280–301. doi: 10.1016/j.bandl.2005.06.001
- Hermosillo, R., Ritterband-Rosenbaum, A., and van Donkelaar, P. (2011). Predicting future sensorimotor states influences current temporal decision making. *J. Neurosci.* 31, 10019–10022. doi: 10.1523/JNEUROSCI.0037-11.2011
- Hirsh, I. J., and Sherrick, C. E. Jr. (1961). Perceived order in different sense modalities. *J. Exp. Psychol.* 62, 423–432. doi: 10.1037/h0045283
- Kayser, C., Petkov, C. I., and Logothetis, N. K. (2008). Visual modulation of neurons in auditory cortex. *Cereb. Cortex* 18, 1560–1574. doi: 10.1093/cercor/bhm187
- Keetels, M., and Vroomen, J. (2005). The role of spatial disparity and hemifields in audio-visual temporal order judgments. *Exp. Brain Res.* 167, 635–640. doi: 10.1007/s00221-005-0067-1
- Keetels, M., and Vroomen, J. (2008). Temporal recalibration to tactile-visual asynchronous stimuli. *Neurosci. Lett.* 430, 130–134. doi: 10.1016/j.neulet.2007.10.044
- Kitagawa, N., Kato, M., and Kashino, M. (2009). “Assessing the effect of voluntary action on sensitivity to temporal asynchrony between auditory and somatosensory events,” in *Poster Presented at the 10th International Multisensory Research Forum*, New York.
- Kitagawa, N., Zampini, M., and Spence, C. (2005). Audio-tactile interactions in near and far space. *Exp. Brain Res.* 166, 528–537. doi: 10.1007/s00221-005-2393-8
- Kwon, J., Ogawa, K., and Miyake, Y. (2014). The effect of visual apparent motion on audiovisual simultaneity. *PLoS ONE* 9:e110224. doi: 10.1371/journal.pone.0110224
- Lewald, J., and Guski, R. (2003). Cross-modal perceptual integration of spatially and temporally disparate auditory and visual stimuli. *Cogn. Brain Res.* 16, 468–478. doi: 10.1016/S0926-6410(03)00074-0
- London, B. M., and Miller, L. E. (2013). Responses of somatosensory area 2 neurons to actively and passively generated limb movements. *J. Neurophysiol.* 109, 1505–1513. doi: 10.1152/jn.00372.2012
- Miller, J., and Schwarz, W. (2006). Dissociations between reaction times and temporal order judgments: a diffusion model approach. *J. Exp. Psychol. Hum. Percept. Perform.* 32, 394–412. doi: 10.1037/0096-1523.32.2.394
- Mitrani, L., Shekardjiski, S., and Yakimoff, N. (1986). Mechanisms and asymmetries in visual perception of simultaneity and temporal order. *Biol. Cybern.* 54, 159–165. doi: 10.1007/BF00356854
- Morrone, M. C., Ross, J., and Burr, D. (2005). Saccadic eye movements cause compression of time as well as space. *Nat. Neurosci.* 8, 950–954. doi: 10.1038/nn1488
- Nijhawan, R. (2002). Neural delays, visual motion and the flash-lag effect. *Trends Cogn. Sci.* 6, 387–393. doi: 10.1016/S1364-6613(02)01963-0
- Nishi, A., Yokoyama, M., Ogawa, K., Ogata, T., Nozawa, T., and Miyake, Y. (2014). Effects of voluntary movements on audio-tactile temporal order judgment. *IEICE Trans. Inf. Syst.* 6, 1567–1573. doi: 10.1109/SII.2011.6147525
- Nishikawa, N., Yasushi, S., Makoto, W., Nobutaka, H., and Kitazawa, S. (2015). Effects of aging and idiopathic Parkinson's disease on tactile temporal order judgment. *PLoS ONE* 10:e0118331. doi: 10.1371/journal.pone.0118331
- Petrini, K., Russell, M., and Pollick, F. (2009). When knowing can replace seeing in audiovisual integration of actions. *Cognition* 110, 432–439. doi: 10.1016/j.cognition.2008.11.015
- Rauschecker, J., and Scott, S. (2009). Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat. Neurosci.* 12, 718–724. doi: 10.1038/nn.2331
- Sanders, L. D., and Astheimer, L. B. (2008). Temporally selective attention modulates early perceptual processing: event-related potential evidence. *Percept. Psychophys.* 70, 732–742. doi: 10.3758/PP.70.4.732
- Schneider, K. A., and Bavelier, D. (2003). Components of visual prior entry. *Cogn. Psychol.* 47, 333–366. doi: 10.1016/S0010-0285(03)00035-5
- Shi, Z., Hirsch, S., Schneider, W., and Müller, H. (2008). “Influence of visuomotor action on visual-haptic simultaneous perception: a psychophysical study,” in *Proceeding of the Symposium on Haptic Interfaces for Virtual Environments and Teleoperator Systems* (Reno, NE: IEEE), 65–70.
- Shore, D. I., Spence, C., and Klein, R. M. (2001). Visual prior entry. *Psychol. Sci.* 12, 205–212. doi: 10.1111/1467-9280.00337
- Slutsky, D. A., and Recanzone, G. H. (2001). Temporal and spatial dependency of the ventriloquism effect. *Neurol. Rep.* 12, 7–10.
- Spence, C., Shore, D. I., and Klein, R. M. (2001). Multisensory prior entry. *J. Exp. Psychol. Gen.* 130, 799–832. doi: 10.1037/0096-3445.130.4.799
- Takahashi, T., Kansaku, K., Wada, M., Shibuya, S., and Kitazawa, S. (2013). Neural correlates of tactile temporal-order judgment in humans: an fMRI study. *Cereb. Cortex* 23, 1952–1964. doi: 10.1093/cercor/bhs179
- Tanji, J., and Mushiake, H. (1996). Comparison of neuronal activity in the supplementary motor area and primary motor cortex. *Cogn. Brain Res.* 3, 143–150. doi: 10.1016/0926-6410(95)00039-9
- Tian, X., and Poeppel, D. (2010). Mental imagery of speech and movement implicates the dynamics of internal forward models. *Front. Psychol.* 1:166. doi: 10.3389/fpsyg.2010.00166
- Van Eijk, R. L. J., Kohlrausch, A., Juola, J. F., and van de Par, S. (2009). Temporal interval discrimination thresholds depend on perceived synchrony for audio-visual stimulus pairs. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 1254–1263. doi: 10.1037/a0014254

- Vogels, I. M. (2004). Detection of temporal delays in visual-haptic interfaces. *Hum. Factors* 46, 118–134. doi: 10.1518/hfes.46.1.118.30394
- Vroomen, J., and Stekelenburg, J. J. (2010). Visual anticipatory information modulates multisensory interactions of artificial audiovisual stimuli. *J. Cogn. Neurosci.* 22, 1583–1596. doi: 10.1162/jocn.2009.21308
- Waszak, F., Cardoso-Leite, P., and Hughes, G. (2012). Action effect anticipation: neurophysiological basis and functional consequences. *Neurosci. Biobehav. Rev.* 36, 943–959. doi: 10.1016/j.neubiorev.2011.11.004
- Weber, D. J., London, B. M., Hokanson, J. A., Ayers, C. A., Gaunt, R. A., Torres, R. R., et al. (2011). Limb-state information encoded by peripheral and central somatosensory neurons: implications for an afferent interface. *IEEE Trans. Neural Syst. Rehabil. Eng.* 19, 501–513. doi: 10.1109/TNSRE.2011.2163145
- Wenke, D., and Haggard, P. (2009). How voluntary actions modulate time perception. *Exp. Brain Res.* 196, 311–318. doi: 10.1007/s00221-009-1848-8
- Yarrow, K., Haggard, P., Heal, R., Brown, P., and Rothwell, J. C. (2001). Illusory perceptions of space and time preserve cross-saccadic perceptual continuity. *Nature* 414, 302–305. doi: 10.1038/35104551
- Yokoyama, M., Yoshida, S., Ora, H., and Miyake, Y. (2009). “Effect of voluntary motion in inter-modal simultaneity judgment,” in *Proceedings of Human Interface Symposium*, Tokyo, 877–882.
- Zampini, M., Brown, T., Shore, D. I., Maravita, A., Röder, B., and Spence, C. (2005a). Audiotactile temporal order judgments. *Acta Psychol.* 118, 277–291. doi: 10.1016/j.actpsy.2004.10.017
- Zampini, M., Guest, S., Shore, D. I., and Spence, C. (2005b). Audio-visual simultaneity judgments. *Percept. Psychophys.* 67, 531–544. doi: 10.3758/BF03193329
- Zampini, M., Shore, D. I., and Spence, C. (2005c). Audiovisual prior entry. *Neurosci. Lett.* 381, 217–222. doi: 10.1016/j.neulet.2005.01.085
- Zampini, M., Shore, D. I., and Spence, C. (2003). Audiovisual temporal order judgments. *Exp. Brain Res.* 152, 198–210. doi: 10.1007/s00221-003-1536-z

**Conflict of Interest Statement:** 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.

Copyright © 2015 Hao, Ogata, Ogawa, Kwon and Miyake. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





# Wearing weighted backpack dilates subjective visual duration: the role of functional linkage between weight experience and visual timing

Lina Jia<sup>1\*</sup>, Zhuanghua Shi<sup>2</sup> and Wenfeng Feng<sup>3\*</sup>

<sup>1</sup> Department of Education, School of Humanities, Jiangnan University, Wuxi, China, <sup>2</sup> Department of Psychology, Ludwig-Maximilians-Universität München, Munich, Germany, <sup>3</sup> Department of Psychology, School of Education, Soochow University, Suzhou, China

## OPEN ACCESS

### Edited by:

Yan Bao,  
Peking University, China

### Reviewed by:

Lihua Mao,  
Peking University, China  
Aneta Szymaszek,  
Nencki Institute of Experimental  
Biology, Poland

### \*Correspondence:

Lina Jia,  
Department of Education, School  
of Humanities, Jiangnan University,  
Li Lake Street 1800, Wuxi 214122,  
China  
jialina09@gmail.com;  
Wenfeng Feng,  
Department of Psychology, School  
of Education, Soochow University,  
Ren-Ai Road 199, Suzhou 215123,  
China  
fengwfly@gmail.com

### Specialty section:

This article was submitted to  
Perception Science,  
a section of the journal  
Frontiers in Psychology

**Received:** 01 July 2015

**Accepted:** 26 August 2015

**Published:** 08 September 2015

### Citation:

Jia L, Shi Z and Feng W (2015)  
Wearing weighted backpack dilates  
subjective visual duration: the role  
of functional linkage between weight  
experience and visual timing.  
Front. Psychol. 6:1373.  
doi: 10.3389/fpsyg.2015.01373

Bodily state plays a critical role in our perception. In the present study, we asked the question whether and how bodily experience of weights influences time perception. Participants judged durations of a picture (a backpack or a trolley bag) presented on the screen, while wearing different weight backpacks or without backpack. The results showed that the subjective duration of the backpack picture was dilated when participants wore a medium weighted backpack relative to an empty backpack or without backpack, regardless of identity (e.g., color) of the visual backpack. However, the duration dilation was not manifested for the picture of trolley bag. These findings suggest that weight experience modulates visual duration estimation through the linkage between the wore backpack and to-be-estimated visual target. The congruent action affordance between the wore backpack and visual inputs plays a critical role in the functional linkage between inner experience and time perception. We interpreted our findings within the framework of embodied time perception.

**Keywords:** duration estimation, weight, action affordance, bodily states, embodiment

## Introduction

Our cognition and perception are grounded in bodily state as well as its interaction with environment (Clark, 1999; Barsalou, 2003). For example, when observers wear a heavy backpack, the geographical slant is likely to be overestimated both in real and virtual hills (Proffitt et al., 1995; Witt and Proffitt, 2007). Similar effects of weight experience have been shown in judgments of spatial distances and monetary values (Witt et al., 2004; Jostmann et al., 2009). When participants threw a heavy ball, the subjective distance was biased by the ball that they threw (Witt et al., 2004). It has been argued (Proffitt, 2006) that such distorted perception reflects the physical energetic costs associated with action plans, as heavy objects, compared to light objects, require more physical strength to act, which is in line with the framework of embodied cognition (Clark, 1999; Barsalou, 2008) that perception, body, and action are tightly linked together.

Not only spatial perception, time perception can also be better understood within the framework of embodiment (Clark, 1999; Droit-Volet et al., 2013; Wittmann, 2013; Maniadakis et al., 2014). Studies have demonstrated that bodily states markedly influence time perception (Yarrow et al., 2001; Droit-Volet and Gil, 2009; Hagura et al., 2012; Shi et al., 2013; Jia et al., 2015). For example, external trains of clicks and intake of drugs (e.g., amphetamine) can change bodily arousal levels,

leading to distortions of perceived durations (Maricq et al., 1981; Penton-Voak et al., 1996). Similarly, studies have shown pictures that are more arousing are often perceived longer than low arousal ones (Droit-Volet and Gil, 2009). Voluntary actions or action preparation can also adjust bodily states, affecting subjective time (Effron et al., 2006; Hagura et al., 2012; Maniadakis et al., 2014; Jia et al., 2015). Noted, action-based bodily regulation may overwrite potential influences of affective stimuli on subjective time. For instance, when participants could freely imitate high-arousal facial expressions presented on the screen, the durations of presented angry and happy faces were often overestimated. But such subjective duration expansion diminished when their imitation of the facial expressions was inhibited by holding a pen between their lips (Effron et al., 2006). A recent study (Jia et al., 2015) has also demonstrated that possibility of stimulus-response interaction could change perceived duration of a tactile stimulus.

To explain interactions between subjective time and interoceptive (bodily) states, Craig (2009) proposed awareness theory based on brain imaging studies. According to this theory, the anterior insula cortex unifies meta-representations of homeostatic feeling states that produce a cinemascopic 'image' of sentient self across time, and subsequently subjective time is estimated through these moments (Craig, 2009; Wittmann et al., 2010; Wittmann, 2013). When a stimulus is related to the survival of body self (e.g., an approaching object toward the observer; see Jia et al., 2015), the inner sentient moments run fast, and subsequently its duration is overestimated. Several recent studies have provided the evidence of this claim (Wittmann et al., 2010; Pollatos et al., 2014; Jia et al., 2015). For instance, the awareness of bodily states influences duration judgments of emotional films (Pollatos et al., 2014). When watching film clips, one group were told to notice their bodily states, whereas the other group were asked to pay attention to the details of film clips to answer several questions later. Afterward, participants recalled the duration of film clips. The results showed that attending to bodily states increased the effects of emotional states on duration judgment compared to attending to clips.

However, the body-related events (e.g., action and emotional) are often accompanied with the changes of arousal, indicated by physiological body response (e.g., increase of skin conductance response and contraction of muscles for threats; Bradley et al., 2001). In other words, bodily states and arousal are hard to separate. In addition, these salient events might capture attention (Vuilleumier, 2005). Another two classic accounts of time perception, the general arousal account and the 'attention-gate' theory, can also partially explain time distortions of body-related events using the internal clock model (Gibbon et al., 1984; Gibbon and Church, 1990). According to the internal clock model, the internal clock consists of a pacemaker, a switch, and an accumulator. The switch is located between the pacemaker and the accumulator. When the switch closes, the temporal pulses emitted by the pacemaker are transmitted to the accumulator where the number of pulses decides the length of subjective duration; when the switch opens, the accumulation process stops. Some body-related stimuli would increase the arousal, according to the arousal account, speeding up the

pacemaker to emit pulses, and resulting in duration dilation (Hodinott-Hill et al., 2002; Droit-Volet et al., 2004; Nather et al., 2011). By contrast, the 'attentional-gate' theory (Block and Zakay, 1997; Zakay and Block, 1997) proposed that attention resources are divided between temporal processing of the clock and non-temporal processing. If a body-related event engages more attention, less attention would be allocated to the timing process in the clock, inducing a loss of some temporal pulses due to the 'flickering' open and closed states of the switch. Consequently, duration is underestimated. These three accounts highlight the importance of self-reference, arousal, and attention factors on duration judgment, respectively. The question of which factors play what critical roles in timing has been hot debated recently (Droit-Volet and Meck, 2007; Droit-Volet and Gil, 2009; Maniadakis et al., 2014). Noted, the self-referential process is often coupled with the change of arousal, with the former emphasizing the interaction between the to-be-estimated stimulus and the observer (embodiment). They could commonly contribute to duration distortions in some body-related contexts, although the self-reference and the sensorimotor states seem to play a more important role than affective states (Effron et al., 2006; Nather et al., 2011; Pollatos et al., 2014).

Previous studies concerning embodied timing interpreted that changes of bodily states (e.g., implicit action) caused by target stimuli are critical for duration distortion of the target stimuli (Droit-Volet and Gil, 2009; Wittmann et al., 2010; Shi et al., 2012; Maniadakis et al., 2014). In most cases, the target stimuli and changes of bodily states have some causal relationship, or at least are highly relevant. However, it is unclear whether a functional linkage between the target stimuli and bodily states is necessary for subjective time distortion, or just the change of bodily states already distorts subjective duration. Investigation of such question would provide a new view of interactions between bodily states and timing. One approach is to examine whether and when the duration of a neutral visual stimulus would be distorted in the context of some specific bodily states induced by nonvisual sources, such as weight experience. It has been suggested that weight experience could change bodily states (Proffitt, 2006). For instance, wearing a heavy backpack requires our bodies to afford with more physical efforts relative to wearing a light backpack, and thus different pressure states in the sensory-motor loop might influence temporal judgment by speeding up internal sentient moments (or / and the clock) of weight experience (Craig, 2009; Nather et al., 2011). Given that the neutral visual stimulus is irrelevant to the change of bodily states activated by the weight experience, can the weight experience still impact on visual time judgments in general? If this is the case, it will suggest that the weight experience affects timing by mediating arousal. Alternative, influences of weight experience on visual duration judgments may require some functional linkages, such as by similar action affordance between weight experience and visual target stimulus. According to the theory of affordance (Gibson, 1979), different objects in the environment have different affordances for manipulation. For example, hammer usually affords hitting, knife cutting, and backpack wearing. In line with this view, neurophysiological studies have revealed that even observing the static picture of a

manipulable object (e.g., tools) could activate the premotor and parietal motor areas (Chao and Martin, 2000; Grezes and Decety, 2002; Kiefer et al., 2011). The affordance offers the possibility of the linkage between external stimulus and bodily states, which might affect perception and cognition. Studies have shown that recognition of a pictorial object was affected by another pictorial object through the congruent action affordance (Helbig et al., 2006; Kiefer et al., 2011). Based on similar reasoning, we hypothesize that the linkage established by the congruent affordance between weight experience ('wearing' behavior) and visual target containing 'wearing' affordance (e.g., backpack picture) might be critical for duration distortion of the visual target. Note that arousal and functional linkage are not mutually exclusive, and both can affect time judgment at the same time. Very heavy weight experience may cause great arousing, which may expand subjective duration in general (Gibbon et al., 1984; Droit-Volet and Gil, 2009). Here, we were most interested in whether the functional linkage mediates weight experience and time judgments, thus we only used medium weight in the study.

The present study was designed to investigate whether and how weighted experience, wearing a 5.7 kg backpack, influences visual duration judgments. In particular, whether congruent action affordance between weight experience and visual target plays a key role in subjective time distortion. Participants were asked to judge the duration of computer-presented pictures, either a backpack or trolley bag, while wearing a real weighted or empty backpack. The function of backpack is 'wearing', whereas the function of trolley bag is 'pulling'. Thus, a backpack picture with affordance of 'wearing', regardless of its feature (e.g., color, style), might activate its functional linkage to weight experience through the congruent affordance. Then the weighted experience induced by the wore weighted backpack, associated with more energy costs, might dilate subjective duration of the backpack picture via such functional linkage, whereas the 'pulling' trolley bag is incongruent in affordance with the 'wearing' backpacks, such that the weight experience may have little influence on duration judgments of the visual trolley bag. Alternatively, the general arousal account (Gibbon et al., 1984; Gibbon and Church, 1990) would predict that the duration distortions induced by the wore backpack, if any, would be similar for both backpack and trolley bag pictures. Similarly, the 'attentional-gate' theory (Zakay and Block, 1997) would predict underestimated durations, if any, for both the backpack and the trolley bag pictures. This is because if attention is distracted by the wore backpack during the time estimation, less attention for the visual timing task would lead to underestimation. To disassociate these alternative accounts, we conducted three experiments. Experiment 1 compared visual duration estimations of the backpack picture among the conditions of wearing weighted backpack, empty backpack, and no backpack conditions. The backpack depicted in a picture was the same to the wearing one. In Experiment 2, we changed the identity of the visual backpack picture, but remained the same congruent 'wearing' affordance. In Experiment 3, we changed the backpack picture to a trolley bag picture, which has different action affordance meanings ('pulling') from the wore backpack ('wearing').

## Materials and Methods

### Participants

Fifty-five students from Jiangnan University took part in the experiments (18, 19, and 18 in Experiments 1, 2, and 3, respectively; 37 female; mean age = 20.7,  $SD = 2.7$ ). The numbers of females were 11, 12, and 14 in Experiments 1, 2, and 3, respectively. All participants had normal or corrected-to-normal visual acuity and no somatosensory disorders. All participants were naive to the purpose of experiments. The experiments were approved by the ethics committee of Jiangnan University. Informed consent in accordance with the Declaration of Helsinki was obtained from each participant before the start of the experiment.

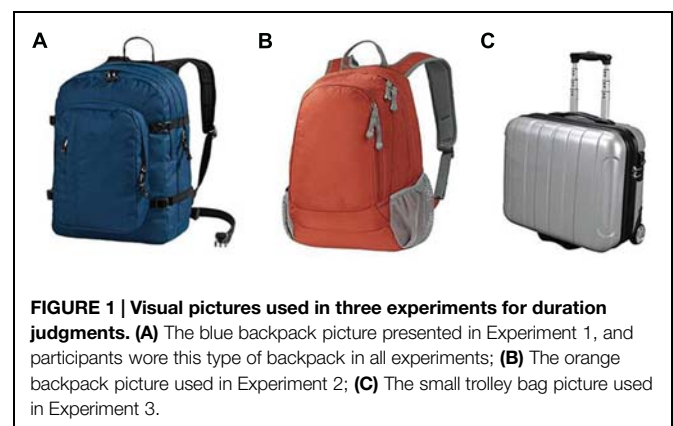
### Stimuli and Apparatus

The experiments were conducted in an isolated cabin with dim lit environment. Visual stimuli were presented on a 21-inch CRT monitor with a refresh rate of 100 Hz. Visual stimuli consisted of the following pictures: blue and orange backpacks (12 cm × 9 cm), small gray business trolley bag (10 cm × 10 cm, see **Figure 1**). Participants were asked to keep standing and holding a light response box during blocks. The viewing distance was kept at 57 cm. Visual stimuli presentation was controlled by Matlab program using Psychophysics Toolbox (Brainard, 1997).

In each trial, the to-be-estimated visual duration was the exposure time of a picture, which could be a blue backpack (Experiment 1), an orange backpack (Experiment 2), or a gray trolley bag (Experiment 3). During all experiments, participants wore a blue backpack (44 cm × 32 cm × 35cm) depicted in **Figure 1A**. Prior to the experiment, participants were told that the weights of the blue backpack (**Figure 1A**), orange backpack (**Figure 1B**), or small trolley bag (**Figure 1C**) in pictures were the same as the wore blue backpack (weighted or empty).

### Experimental Procedure

A classic temporal bisection task was used in the experiments. Participants were first trained to discriminate two visual anchor durations: a short one (200 ms) and a long one (600 ms). The anchor stimulus was a white rectangle (12 cm × 9 cm



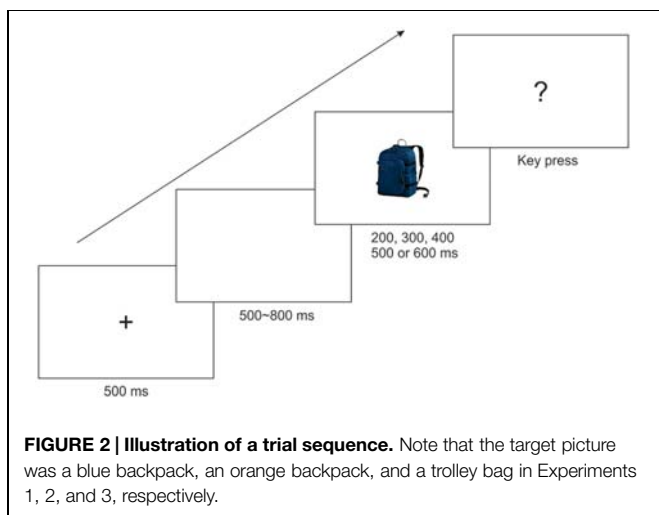
**FIGURE 1 | Visual pictures used in three experiments for duration judgments. (A)** The blue backpack picture presented in Experiment 1, and participants wore this type of backpack in all experiments; **(B)** The orange backpack picture used in Experiment 2; **(C)** The small trolley bag picture used in Experiment 3.

for Experiments 1 and 2; 10 cm × 10 cm for Experiment 3), same size as the pictures used in the experiments. The training session ended when participants reached 100% accuracy of discrimination for consecutive 20 trials.

In the subsequent test session, illustrated in **Figure 2**, each trial started with a fixation cross for 500 ms, followed by a blank display randomly for 500~800 ms. Then a target picture (backpack in Experiments 1 and 2, trolley bag in Experiment 3) was presented for a given probe duration, randomly selected from 200, 300, 400, 500, or 600 ms. After the picture presentation, a question mark was shown to prompt for a response. Participants had to judge whether the duration of the picture was closer to the short anchor (200 ms) or the long anchor (600 ms) as accurately as possible by pressing the left or right key on the response box, respectively. The inter-trial interval (ITI) varied randomly from 1000 to 1500 ms.

The test session consisted of three conditions of wearing weights block-wisely: the weighted backpack (5.7 kg), empty backpack (0.7 kg), and no backpack (baseline) conditions. Each weight condition was repeated twice, and randomly intermixed with the other conditions. Within each block, five probe durations were repeated randomly for 10 times, yielding 50 trials per block. Thus, the test session consisted of 300 trials. To refresh participants about the short and long anchors, each of the two anchors was presented for five times at the beginning of each block. Participants took a rest about 2 min by taking off the backpack between blocks. The length of the test session was around 50 min.

After the test session, participants were asked to rate the valence and arousal using the paper sheet of the affective-rating Self-Assessment-Manikin (SAM) in order to compare the arousal levels among three conditions of wearing weights. The SAM evaluation is 9-point scales rating, ranging from sad to pleasant for the 'valence' and from calm to activated for the 'arousal' (Bradley and Lang, 1994). To make sure that participants understood the meanings of 9 points on valence and arousal scales, respectively, they were presented with the detailed instruction before their evaluation.

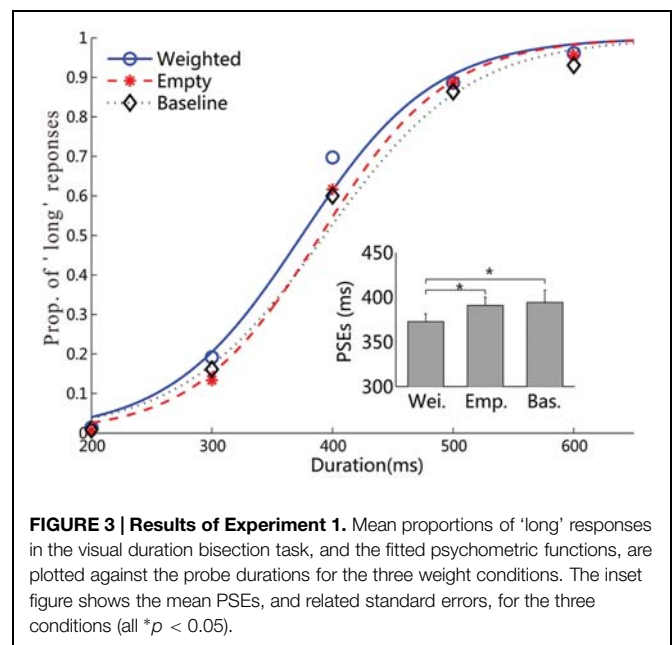


## Results

The proportions of 'long' responses for the five probe durations were calculated and fitted by a logistic function for each participant at each weight condition. The points of subjective equality (PSEs) of the temporal bisection were then estimated corresponding to the duration at the 50th percentile of the fitted curves. To measure the sensitivity of duration judgments, the just-noticeable differences (JNDs) were estimated by taking half the difference in durations between the 25th and 75th percentiles (see detailed method in Shi et al., 2012). Repeated-measures ANOVAs with wearing weight as factor were conducted separately on the PSEs and JNDs in all experiments, and then further LSD contrast tests were performed to see the significant differences among conditions of wearing weights. Similar ANOVAs were applied for subjective arousal ratings.

### Duration Judgment

Experiment 1 examined the influences of wearing a backpack on the duration judgment of the same backpack picture. **Figure 3** shows the psychometric curves of the visual-duration bisection task for the weighted backpack, empty backpack, and baseline conditions, respectively. The mean PSEs ( $\pm$ SE) were  $373 \pm 9$ ,  $391 \pm 9$ , and  $394 \pm 13$  ms for the weighted backpack, empty backpack, and baseline conditions (**Table 1**). Repeated measures ANOVA revealed a significant influence of wearing weights on the visual duration judgment,  $F(2,34) = 3.66$ ,  $p < 0.05$ ,  $\eta_p^2 = 0.18$ . The further post-hoc contrast tests showed significant differences in PSEs between the weighted and empty backpack conditions (difference: 18 ms,  $p < 0.05$ ), and between the weighted backpack and baseline conditions (difference: 21 ms,  $p < 0.05$ ), but not between the empty backpack and baseline conditions ( $p = 0.70$ ). The JNDs ( $\pm$ SE) were  $53 \pm 6$ ,  $50 \pm 4$ , and  $55 \pm 3$  ms for the weighted backpack, empty backpack, and baseline conditions





**TABLE 1 | Mean of points of subjective equality (PSEs) and just-noticeable differences (JNDs) for three weight conditions across all experiments (ms).**

	PSE(±SE)			JND(±SE)		
	Weighted	Empty	Baseline	Weighted	Empty	Baseline
Experiment 1	373(9)	391(9)	394(13)	53(6)	50(4)	55(3)
Experiment 2	388(11)	412(10)	404(13)	53(4)	62(5)	54(3)
Experiment 3	398(12)	395(13)	401(12)	55(3)	58(4)	51(4)

(Table 1). A repeated-measures ANOVA failed to show any significant difference on JNDs among these three conditions,  $F(2,34) = 0.69$ ,  $p = 0.51$ ,  $\eta_p^2 = 0.04$ .

Experiment 2 changed the identity of the backpack picture, yielding similar results as those of Experiment 1 (Figure 4). The mean PSEs (±SE) were  $388 \pm 11$ ,  $412 \pm 10$ , and  $404 \pm 13$  ms for the weighted, empty backpacks and baseline conditions, respectively (Table 1). The ANOVA revealed that the influence of the feeling of weight on visual duration judgments was significant,  $F(2,36) = 3.5$ ,  $p < 0.05$ ,  $\eta_p^2 = 0.16$ . The post-hoc contrasts showed significant differences in PSEs between the weighted and empty backpack conditions, the weighted and baseline conditions, respectively (differences: 24 and 16 ms, both  $p < 0.05$ ), but no significant difference between the empty backpack and baseline conditions ( $p = 0.44$ ). A further ANOVA on discrimination sensitivity (JNDs) showed a marginal significance among three conditions,  $F(2,36) = 3.3$ ,  $p = 0.05$ ,  $\eta_p^2 = 0.16$ . Further contrast tests indicated that the JND in the weighted backpack condition was significantly lower than that in the empty backpack condition ( $p < 0.05$ ), while no differences were shown in other comparison conditions (weighted backpack vs. baseline:  $p = 0.83$ ; empty backpack vs. baseline:  $p = 0.07$ ).

Experiment 3, on the other hand, revealed different outcomes (Figure 5). The mean PSEs (±SE) were in similar magnitudes for the three conditions:  $398 \pm 12$ ,  $395 \pm 13$ , and  $401 \pm 12$  ms for the weighted, empty backpacks and baseline, respectively (Table 1), and failed to reveal any main effect of perceiving weight on the

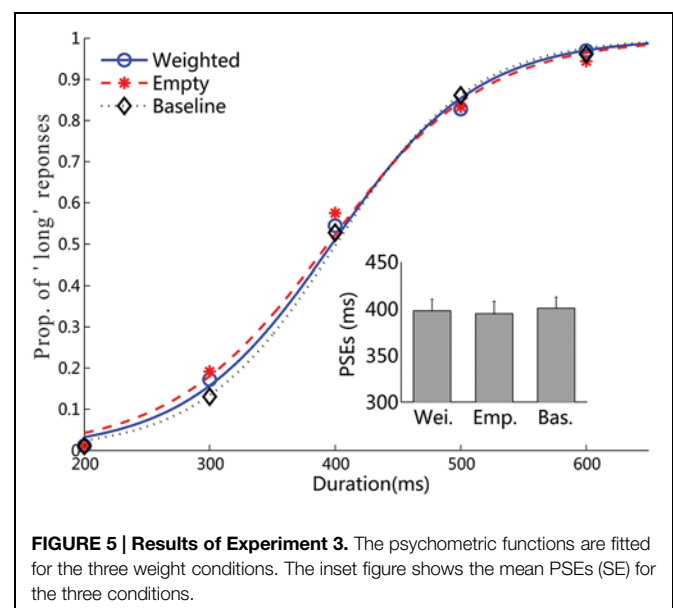
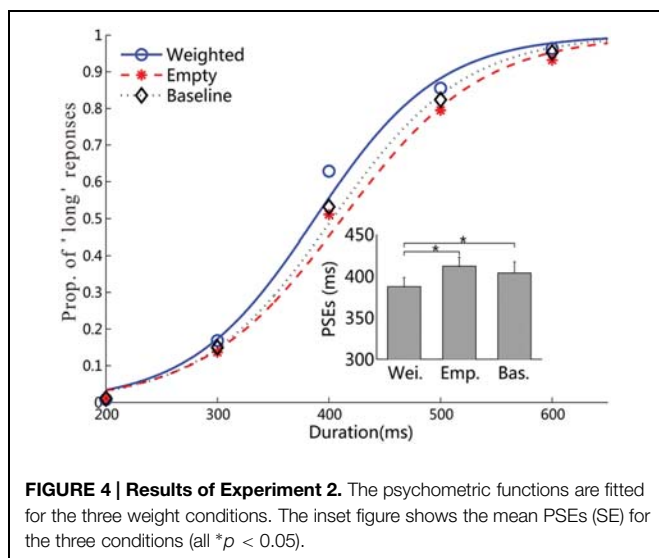
visual duration judgment,  $F(2,34) = 0.28$ ,  $p = 0.76$ ,  $\eta_p^2 = 0.02$ . Similar to the previous two experiments, the JNDs (±SE) also failed to show any significant difference among three conditions,  $F(2,34) = 1.23$ ,  $p = 0.31$ ,  $\eta_p^2 = 0.07$ .

### Assessment of Arousal

Given that subjective ratings of arousal were similar across three experiments, we collapsed arousal ratings across three experiments for the weighted backpack, empty backpack, and baseline conditions. The total results showed that the subjective ratings of arousal significantly differed among the three conditions (Greenhouse–Geisser Correction:  $F(2,108) = 5.14$ ,  $p < 0.05$ ,  $\eta_p^2 = 0.09$ ). The further contrast tests showed that both weighted (mean 4.95) and empty (mean 4.72) backpacks were rated to be more arousing than the baseline (mean 4.53) (both  $p < 0.05$ ), but there was no evidence of significant difference between the weighted and empty backpack conditions ( $p = 0.12$ ).

### Discussion

The present study examined how wearing a medium weighted backpack modulated visual duration judgments. We found that wearing a weighted backpack (5.7 kg) lengthened subjective duration of a backpack picture, regardless of the identity of the backpack. In contrast, weight experience failed to impact



duration judgments of a trolley bag picture. The findings suggest that the effect of weight experience on visual duration judgments depends on a functional link between weight experience and the to-be-estimated picture.

Our findings of differential impacts of weight experiences on visual duration judgments could hardly be explained by the arousal-based account (Gibbon et al., 1984) or the attentional-gating theory (Zakay and Block, 1997). SAM evaluation showed that the rated arousal levels for the weighted and empty backpacks were higher than baseline condition. According to the arousal account, visual duration should be expanded, if any, across all experiments for wearing backpack compared to not wearing backpack. However, the duration expansion effect was only revealed for the weighted backpack condition in Experiments 1 and 2 where the visual targets were backpack pictures, but not in Experiment 3 where the visual target was a trolley bag. Alternatively, attentional-gating theory would predict that attention shifts away from the duration judgment task, if any, to the weight experience, the visual duration would be underestimated, not overestimated. However, such underestimation was not observed in our experiments. Moreover, both arousal and attention accounts would predict reduced temporal sensitivity of temporal bisection in the wore backpack condition compared to the baseline condition, which was not the case in our study as we failed to find their significant differences in JNDs. It should be pointed out, we do not argue that attention and arousal states cannot affect duration judgments (in fact, they do significantly influence duration judgments shown in other studies), rather we suggest merely using attention and arousal states cannot explain the present findings.

Alternatively, our findings can be better explained by the awareness theory based on the embodiment framework (Craig, 2009; Wittmann, 2013; Maniadakis et al., 2014), according to which time perception is an accumulation process of self-related moments (Craig, 2009; Wittmann et al., 2010). In line with this view, recent studies have shown the modulation of near-body arousing stimuli on duration judgment (Effron et al., 2006; Wittmann and van Wassenhove, 2009; Shi et al., 2012; Jia et al., 2015). The bodily experience initiated by the to-be-estimated stimuli with near-body meaning might speed up inner sentient 'moments', leading to duration dilation (Wittmann et al., 2010; Pollatos et al., 2014). It should be noted that in most previous studies bodily states are directly manipulated by affective stimuli or related actions, which are closely related to duration judgments. The present study, on the other hand, provides the first evidence that the functional linkage between timing task and self-referential process is important for the interactions between visual duration judgments and weight experience. The activity of weight pressure was irrelevant to the to-be-estimated target (here the backpack or trolley bag picture), but they could be automatically linked through congruent action affordance. Specifically, when the visual target was the 'wearable' backpack but not a trolley bag, similar to what they wore in affordance, the inner sentient moments for the weight experience and visual estimation were possibly merged together, biasing the time estimation of the visual input in the weighted backpack condition. By contrast, when the visual input had different action

affordances (e.g., 'pulling' of the trolley bag), the inner sentient moments for the weight experience and visual stimulation were likely to be separated, resulting in no effect of weight experience on visual duration judgments. Similar congruency effect of affordance has been demonstrated in response performance (Chen and Bargh, 1999; Alexopoulos and Ric, 2007). The present study extended the affordance congruency effect to duration judgments.

One might argue, however, the same category ('backpack'), rather than congruent affordance, between weight experience and visual blue or orange backpacks, contributed to the linkage. Both the visual blue and orange backpacks can be categorized as 'backpack', but trolley bag cannot. Thus, the category linkage between visual backpacks and the wore backpack might be proposed to induce the impact of weight on visual timing. Gibson (1979) assumed that the same category (defined by the common features) just means a conceptual 'family resemblance' and does not correspond to the congruent affordance. We believe affordance congruency, rather than same category, provides direct linkage between bodily states and time process. First, it has been shown that similar action affordance, not the same category, modulated task performance (Helbig et al., 2006; Weatherford et al., 2015). For example, recognition of the target object following a prime object was facilitated when two objects had the congruent action affordance, although could be classified differently (e.g., pan-dustpan) (Helbig et al., 2006). Second, objects with congruent affordance elicited common neural activities related to motor (Kiefer et al., 2011; Sim et al., 2015), which provides potential mechanism underlying the interaction between perception and bodily states. On this ground, we believe that the congruent affordance between the visual input and weight experience contributed to duration distortion. Still, physiological measures should be used in future work to identify a neural linkage between weight experience and visual timing through the congruent affordance.

## Conclusion

The present research extends the evidence of embodied timing by revealing that wearing a weighted backpack dilates subjective visual duration through a functional linkage. The congruent action affordance between wearing behavior of weight and visual target is critical for such functional role of weight experience on visual timing. Note that we only applied three types of stimuli in visual modality. Thus, future work should expand stimuli to more general categories, and focus on influences of various types of action linkage between weight experience and duration estimation by using different types of sensory inputs, not limited to visual modality.

## Acknowledgments

This study was supported by grants from the Fundamental Research Funds for the Central Universities of China (JUSRP11580, JUSRP51331B) and Natural Science Foundation of China (NSFC 31400868).

## References

- Alexopoulos, T., and Ric, F. (2007). The evaluation-behavior link: direct and beyond valence. *J. Exp. Soc. Psychol.* 43, 1010–1016. doi: 10.1016/j.jesp.2006.10.017
- Barsalou, L. W. (2003). Situated simulation in the human conceptual system. *Lang. Cog. Proc.* 18, 513–562. doi: 10.1080/01690960344000026
- Barsalou, L. W. (2008). Grounded cognition. *Annu. Rev. Psychol.* 59, 617–645. doi: 10.1146/annurev.psych.59.103006.093639
- Block, R. A., and Zakay, D. (1997). Prospective and retrospective duration judgments: a meta-analytic review. *Psychon. Bull. Rev.* 4, 184–197. doi: 10.3758/BF03209393
- 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: 10.1037/1528-3542.1.3.300
- Bradley, M. M., and Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *J. Behav. Ther. Exp. Psychiatry* 25, 49–59. doi: 10.1016/0005-7916(94)90063-9
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spat. Vis.* 10, 433–436.
- Chao, L. L., and Martin, A. (2000). Representation of manipulable man-made objects in the dorsal stream. *Neuroimage* 12, 478–484. doi: 10.1006/nimg.2000.0635
- Chen, M., and Bargh, J. A. (1999). Consequences of automatic evaluation: immediate behavioral predispositions to approach or avoid the stimulus. *Pers. Soc. Psychol. Bull.* 25, 215–224. doi: 10.1177/0146167299025002007
- Clark, A. (1999). An embodied cognitive science? *Trends Cogn. Sci.* 3, 345–351. doi: 10.1016/S1364-6613(99)01361-3
- Craig, A. D. (2009). Emotional moments across time: a possible neural basis for time perception in the anterior insula. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 364, 1933–1942. doi: 10.1098/rstb.2009.0008
- Droit-Volet, S., Brunot, S., and Niedenthal, P. M. (2004). Perception of the duration of emotional events. *Cogn. Emot.* 18, 849–858. doi: 10.1080/02699930341000194
- Droit-Volet, S., Fayolle, S., Lamotte, M., and Gil, S. (2013). Time, emotion and the embodiment of timing. *Timing Time Percept.* 1, 99–126. doi: 10.1163/22134468-00002004
- Droit-Volet, S., and Gil, S. (2009). The time-emotion paradox. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 364, 1943–1953. doi: 10.1098/rstb.2009.0013
- Droit-Volet, S., and Meck, W. H. (2007). How emotions colour our perception of time. *Trends Cogn. Sci.* 11, 504–513. doi: 10.1016/j.tics.2007.09.008
- Effron, D. A., Niedenthal, P. M., Gil, S., and Droit-Volet, S. (2006). Embodied temporal perception of emotion. *Emotion* 6, 1–9. doi: 10.1037/1528-3542.6.1.1
- Gibbon, J., and Church, R. M. (1990). Representation of time. *Cognition* 37, 23–54. doi: 10.1016/0010-0277(90)90017-E
- Gibbon, J., Church, R. M., and Meck, W. H. (1984). Scalar timing in memory. *Ann. N. Y. Acad. Sci.* 423, 52–77. doi: 10.1111/j.1749-6632.1984.tb23417.x
- Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Boston, MA: Houghton Mifflin.
- Grezes, J., and Decety, J. (2002). Does visual perception of object afford action? Evidence from a neuroimaging study. *Neuropsychologia* 40, 212–222. doi: 10.1016/S0028-3932(01)00089-6
- Hagura, N., Kanai, R., Orgs, G., and Haggard, P. (2012). Ready steady slow: action preparation slows the subjective passage of time. *Proc. Biol. Sci.* 279, 4399–4406. doi: 10.1098/rspb.2012.1339
- Helbig, H. B., Graf, M., and Kiefer, M. (2006). The role of action representations in visual object recognition. *Exp. Brain Res.* 174, 221–228. doi: 10.1007/s00221-006-0443-5
- Hodinott-Hill, I., Thilo, K. V., Cowey, A., and Walsh, V. (2002). Auditory chronostasis: hanging on the telephone. *Curr. Biol.* 12, 1779–1781. doi: 10.1016/S0960-9822(02)01219-8
- Jia, L., Shi, Z., Zang, X., and Müller, H. J. (2015). Watching a real moving object expands tactile duration: role of task-irrelevant action context for subjective time. *Atten. Percept. Psychophys.* [Epub ahead of print].
- Jostmann, N. B., Lakens, D., and Schubert, T. W. (2009). Weight as an embodiment of importance. *Psychol. Sci.* 20, 1169–1174. doi: 10.1111/j.1467-9280.2009.02426.x
- Kiefer, M., Sim, E. J., Helbig, H., and Graf, M. (2011). Tracking the time course of action priming on object recognition: evidence for fast and slow influences of action on perception. *J. Cogn. Neurosci.* 23, 1864–1874. doi: 10.1162/jocn.2010.21543
- Maniadakis, M., Wittmann, M., Droit-Volet, S., and Choe, Y. (2014). Toward embodied artificial cognition: time is on my side. *Front. Neurobot.* 8:25. doi: 10.3389/fnbot.2014.00025
- Maricq, A. V., Roberts, S., and Church, R. M. (1981). Methamphetamine and time estimation. *J. Exp. Psychol. Anim. Behav. Process.* 7, 18–30. doi: 10.1037/0097-7403.7.1.18
- Nather, F. C., Bueno, J. L. O., Bigand, E., and Droit-Volet, S. (2011). Time changes with the embodiment of another's body posture. *PLoS ONE* 6:e19818. doi: 10.1371/journal.pone.0019818
- Penton-Voak, I. S., Edwards, H., Percival, A., and Wearden, J. H. (1996). Speeding up an internal clock in humans? Effects of click trains on subjective duration. *J. Exp. Psychol. Anim. Behav. Process.* 22, 307–320. doi: 10.1037/0097-7403.22.3.307
- Pollatos, O., Laubrock, J., and Wittmann, M. (2014). Interoceptive focus shapes the experience of time. *PLoS ONE* 9:e86934. doi: 10.1371/journal.pone.0086934
- Proffitt, D. R. (2006). Embodied perception and the economy of action. *Perspect. Psychol. Sci.* 1, 110–122. doi: 10.1111/j.1745-6916.2006.00008.x
- Proffitt, D. R., Bhalla, M., Gossweiler, R., and Midgett, J. (1995). Perceiving geographical slant. *Psychon. Bull. Rev.* 2, 409–428. doi: 10.3758/BF03210980
- Shi, Z., Church, R. M., and Meck, W. H. (2013). Bayesian optimization of time perception. *Trends Cogn. Sci.* 17, 556–564. doi: 10.1016/j.tics.2013.09.009
- Shi, Z., Jia, L., and Müller, H. J. (2012). Modulation of tactile duration judgments by emotional pictures. *Front. Integr. Neurosci.* 6:24. doi: 10.3389/fnint.2012.00024
- Sim, E. J., Helbig, H. B., Graf, M., and Kiefer, M. (2015). When action observation facilitates visual perception: activation in visuo-motor areas contributes to object recognition. *Cereb. Cortex* 25, 2907–2918. doi: 10.1093/cercor/bhu087
- 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
- Weatherford, K., Mills, M., Porter, A. M., and Goolkasian, P. (2015). Target categorization with primes that vary in both congruency and sense modality. *Front. Psychol.* 6:20. doi: 10.3389/fpsyg.2015.00020
- Witt, J. K., and Proffitt, D. R. (2007). Perceived slant: a dissociation between perception and action. *Perception* 36, 249–257. doi: 10.1068/p5449
- Witt, J. K., Proffitt, D. R., and Epstein, W. (2004). Perceiving distance: a role of effort and intent. *Perception* 33, 577–590. doi: 10.1068/p5090
- Wittmann, M. (2013). The inner sense of time: how the brain creates a representation of duration. *Nat. Rev. Neurosci.* 14, 217–223. doi: 10.1038/Nrn3452
- Wittmann, M., and van Wassenhove, V. (2009). The experience of time: neural mechanisms and the interplay of emotion, cognition and embodiment. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 364, 1809–1813. doi: 10.1098/rstb.2009.0003
- Wittmann, M., van Wassenhove, V., Craig, A. D., and Paulus, M. P. (2010). The neural substrates of subjective time dilation. *Front. Hum. Neurosci.* 4:2. doi: 10.3389/fnuro.09.002.2010
- Yarrow, K., Haggard, P., Heal, R., Brown, P., and Rothwell, J. C. (2001). Illusory perceptions of space and time preserve cross-saccadic perceptual continuity. *Nature* 414, 302–305. doi: 10.1038/35104551
- Zakay, D., and Block, R. A. (1997). Temporal cognition. *Curr. Dir. Psychol. Sci.* 6, 12–16. doi: 10.1111/1467-8721.ep11512604

**Conflict of Interest Statement:** 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.

Copyright © 2015 Jia, Shi and Feng. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# How Two Brains Make One Synchronized Mind in the Inferior Frontal Cortex: fNIRS-Based Hyperscanning During Cooperative Singing

Naoyuki Osaka<sup>1\*</sup>, Takehiro Minamoto<sup>2</sup>, Ken Yaoi<sup>1</sup>, Miyuki Azuma<sup>2</sup>,  
Yohko Minamoto Shimada<sup>3</sup> and Mariko Osaka<sup>2,4</sup>

<sup>1</sup> Department of Psychology, Graduate School of Letters, Kyoto University, Kyoto, Japan, <sup>2</sup> Department of Psychology, Graduate School of Human Sciences, Osaka University, Suita, Japan, <sup>3</sup> Center for Baby Science, Doshisha University, Kyoto, Japan, <sup>4</sup> Center for Information and Neural Networks, Osaka University, Suita, Japan

## OPEN ACCESS

### Edited by:

Yan Bao,  
Peking University, China

### Reviewed by:

Tilmann H. Sander,  
Physikalisch-Technische  
Bundesanstalt, Germany  
Taiki Ogata,  
The University of Tokyo, Japan

### \*Correspondence:

Naoyuki Osaka  
nosaka@bun.kyoto-u.ac.jp

### Specialty section:

This article was submitted to  
Perception Science,  
a section of the journal  
Frontiers in Psychology

**Received:** 12 August 2015

**Accepted:** 10 November 2015

**Published:** 26 November 2015

### Citation:

Osaka N, Minamoto T, Yaoi K,  
Azuma M, Minamoto Shimada Y and  
Osaka M (2015) How Two Brains  
Make One Synchronized Mind in the  
Inferior Frontal Cortex: fNIRS-Based  
Hyperscanning During Cooperative  
Singing. *Front. Psychol.* 6:1811.  
doi: 10.3389/fpsyg.2015.01811

One form of communication that is common in all cultures is people singing together. Singing together reflects an index of cognitive synchronization and cooperation of human brains. Little is known about the neural synchronization mechanism, however. Here, we examined how two brains make one synchronized behavior using cooperated singing/humming between two people and hyperscanning, a new brain scanning technique. Hyperscanning allowed us to observe dynamic cooperation between interacting participants. We used functional near-infrared spectroscopy (fNIRS) to simultaneously record the brain activity of two people while they cooperatively sang or hummed a song in face-to-face (FtF) or face-to-wall (FtW) conditions. By calculating the inter-brain wavelet transform coherence between two interacting brains, we found a significant increase in the neural synchronization of the left inferior frontal cortex (IFC) for cooperative singing or humming regardless of FtF or FtW compared with singing or humming alone. On the other hand, the right IFC showed an increase in neural synchronization for humming only, possibly due to more dependence on musical processing.

**Keywords:** hyperscanning, fNIRS, cooperation, singing, humming, inferior frontal cortex

## INTRODUCTION

People's daily life experiences testify to the fact that through cooperation with others we can achieve goals that we could not reach otherwise. Studies seeking to identify the responsible brain mechanisms for cooperation have been unable to reveal details about the synchronization of the neural activations (Frith and Frith, 1999). Consequently, most investigations of social interactions have measured brain activities in only one person at a given time and not the dynamic interaction of two brains simultaneously.

Synchronization during social interactions has been reported using different neuroimaging techniques. For example, functional magnetic resonance imaging (fMRI) has been to observe two participants during a simple interaction game (Montague et al., 2002) or neuroeconomics



(King-Casas et al., 2005), and electroencephalography (EEG) for social interactions (Astolfi et al., 2011), card game (Babiloni et al., 2006), instrument playing (Lindenberger et al., 2009), and cooperative prisoner's dilemma games (De Vico Fallani et al., 2010).

Singing together is a form of cooperation seen in all cultures and makes a suitable model to study the neural mechanisms of synchronization (Mithen, 2005). In animal studies, the vocalizations of monkeys often have a synchronized musical nature to them. This property is heard most dramatically in the rhythmic chattering of gelades, which are close cousins of baboon, and the "duet" singing of paired gibbons (Geissmann, 2002). Additionally, a pair of wrens showed cooperation through males and females rapidly alternating singing syllables (Fortune et al., 2011). Even insects like orthoptera have been observed to show activity akin to duet singing (Bailey, 2003). In humans, the neural synchronization of cooperative singing may have evolutionarily adapted to make a bond of affection in order to strongly bind groups of people (Dunbar, 2010).

Singing together is also attributed to the adaptation of "flow" (Csikszentmihalyi, 1990). "Flow" can be defined as the mental state of operation in which people performing an activity are fully immersed in a feeling of energized focus. Musicians and choir experience flow, which allows them to make a harmonized song that could not be made with a single participant.

Using magnetoencephalography (MEG), Gunji et al. found distinct cortical rhythmic changes in response to singing and humming consistent with the motor control related to sound production (Gunji et al., 2007). In the alpha band, the oscillatory changes for singing were most pronounced in the right premotor and bilateral superior parietal areas. They also found a high frequency band in Broca's area when participants imagined they were singing.

Recently, online and simultaneous two-brain scanning of subjects engaged in interactive tasks has become possible. This new approach, hyperscanning, can be performed using several methods of different spatial and temporal resolution, including MEG, electroencephalography (EEG), functional magnetic resonance imaging (fMRI), and functional near-infrared spectroscopy (fNIRS), to examine how two brains dynamically interact to make a synchronized mind. fNIRS indirectly estimates a brain's neuronal activity by measuring concentration variations in oxygenated hemoglobin (oxy-Hb) and deoxygenated hemoglobin (deoxy-Hb), which have different absorption spectra in the surface brain's blood flow during the task performed.

Only few studies have reported the brain dynamics of social interactions using fNIRS to measure two brains simultaneously. Jiang et al. (2012) found a significant increase in neural synchronization in the left inferior frontal cortex (IFC) during face-to-face dialog between partners but not during a non-face-to-face dialog, while Cui et al. (2012) found synchronization in the right superior frontal cortex during cooperative but not competitive video games. Interestingly, language-based cooperative dialog and video-based spatial cooperative games activated the left and right frontal cortex, respectively. The dialogs in Jiang et al. (2012) are likely related to verbal activity

in Broca's area, which is in the left IFC, while the visuo-spatial cooperative tasks in Cui et al. (2012) are likely related to activity in the right middle to superior frontal cortex.

These distinct areas should be relevant to synchronized singing and humming (the act of singing with open- and closed-lips, respectively), since the production of words during singing should engage Broca's area in the left IFC, while the production of melody during humming would be more related to the right IFC and superior frontal cortex, as reported using dichotic listening (Kimura, 1964). To test this theory, we applied fNIRS to investigate the neural synchronization of two cooperative partners when singing and humming. We also examined the neuronal differences between single and pair (cooperatively) synchronized singing and humming under face-to-face and non-face-to-face conditions by simultaneously measuring two brains.

## MATERIALS AND METHODS

### Participants

Thirty adults participated in the singing experiment (15 pairs, mean age of 22 years; eight male pairs and seven female pairs), and twenty-eight adults (14 pairs, mean age of 21 years: nine male pairs and five female pairs) in the humming experiment. The gender of the participant-pairs was controlled with matched age, and participants in each pair were mutually unfamiliar and assumed independent pairs. We obtained written informed consent from all participants, and the experimental protocol was approved by the Osaka University Institutional Review Board. Participants were paid (5000 yen each) for their participation.

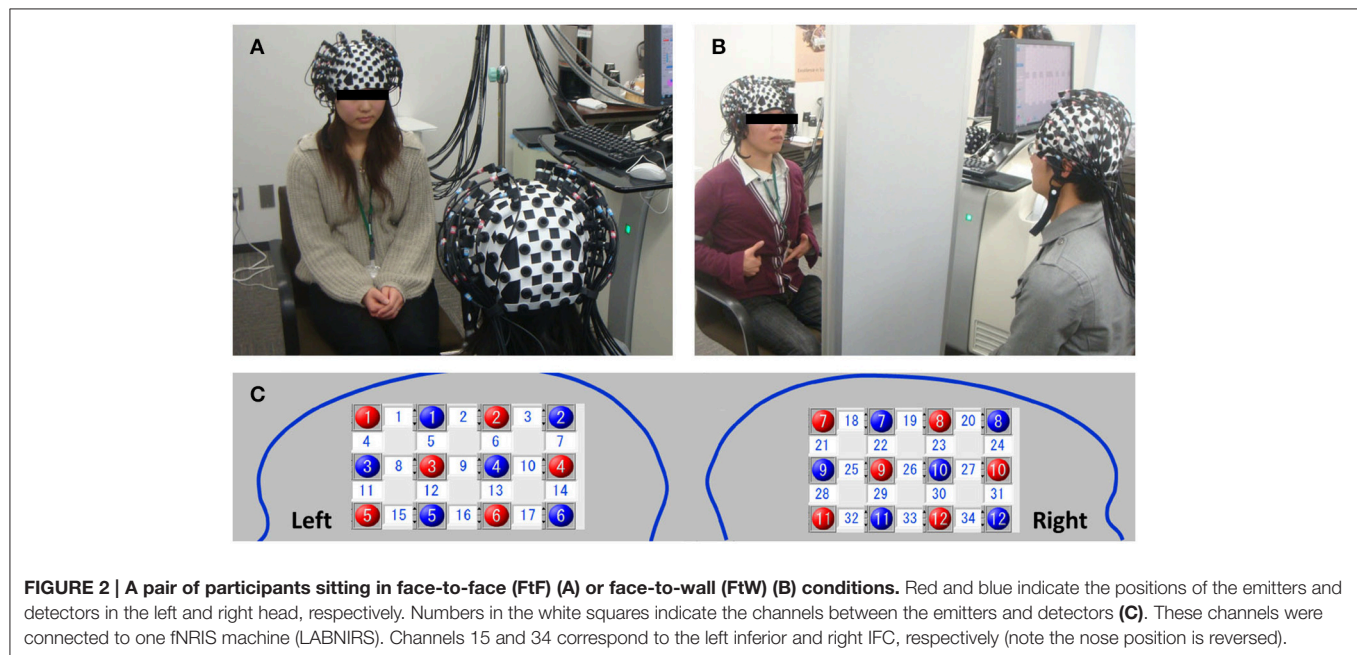
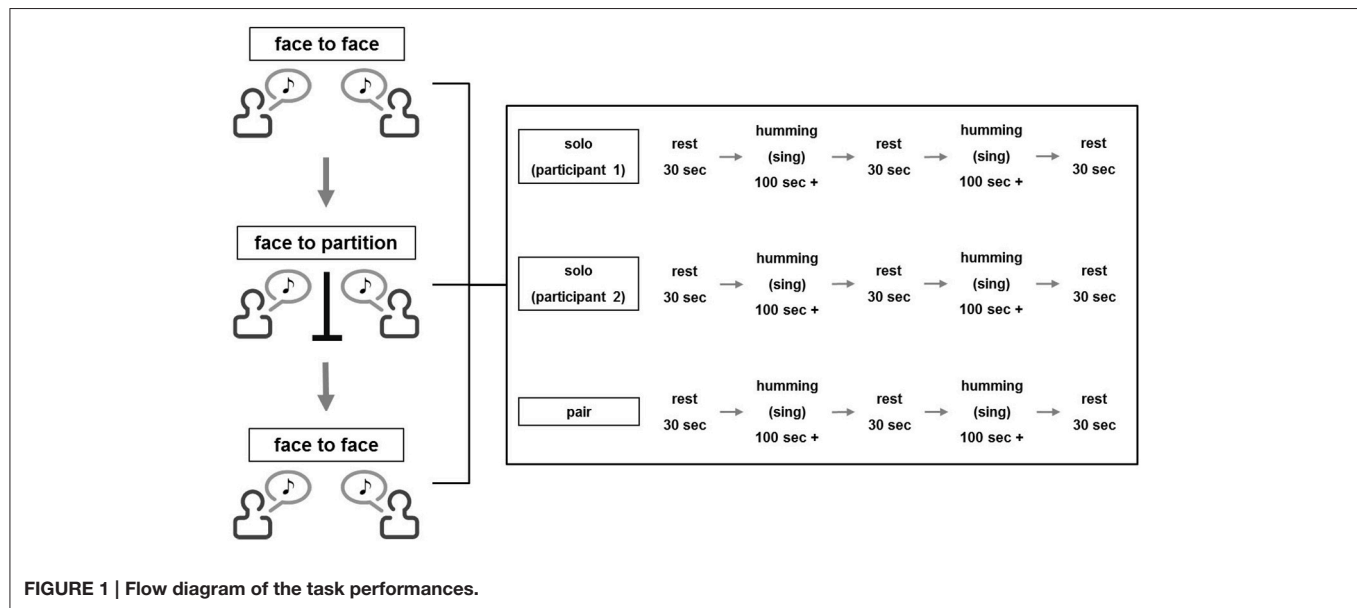
### Stimulus

Three popular Japanese nursery rhymes were selected for the experiments: Under the Spreading Chestnut Tree, School of Killifish, and Sunset with the Evening Glow, since all participants would be familiar with these songs. Participants were instructed to sing or hum a melody part of the song, which lasted for 20–30 s.

### Procedure

In the singing experiment, participants were instructed to sing a song alone, listen to the partner's singing or sing with the partner. The order was counterbalanced across pairs. A time-course of each session is illustrated in **Figure 1**. In accordance with a previous fNIRS study (Cui et al., 2012), a 30-s rest period was given at the beginning of a session. Following the rest, participants sang a song alone or together or listened to the partner sing for about 100 s. When signing, the participant repeated the melody of one of the three songs described above (about 4–5 times) until the 100 s had passed and thereafter stopped singing. When listening, the participant was instructed to actively listen to the partner's singing and gaze at the partner's face. Then, another 30-s rest was given, which was followed by the second 100-s singing/listening interval. One more 30-s break was added at the end of each session. An experimenter measured the time with a stopwatch and instructed the beginning and end of each stage.

Participants performed three experimental sets, where each set consisted of the three conditions (single singing, cooperative



singing, and listening). In the first set, two participants faced each other and performed the set while gazing at the partner's face (first face-to-face condition; FtF). In the second set, an opaque partition was placed between the participants (face-to-wall condition; FtW). In the third set, the partition was removed (second FtF condition). The same song was sung in each set, and the order of the songs was counterbalanced across pairs.

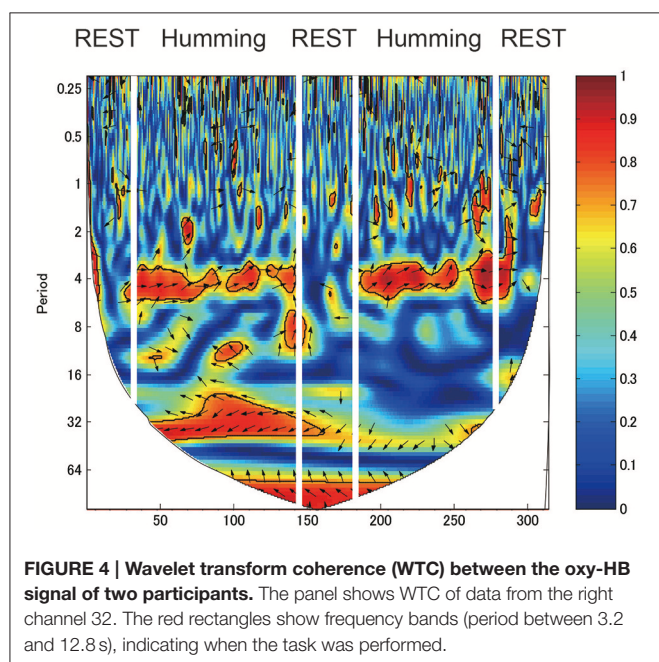
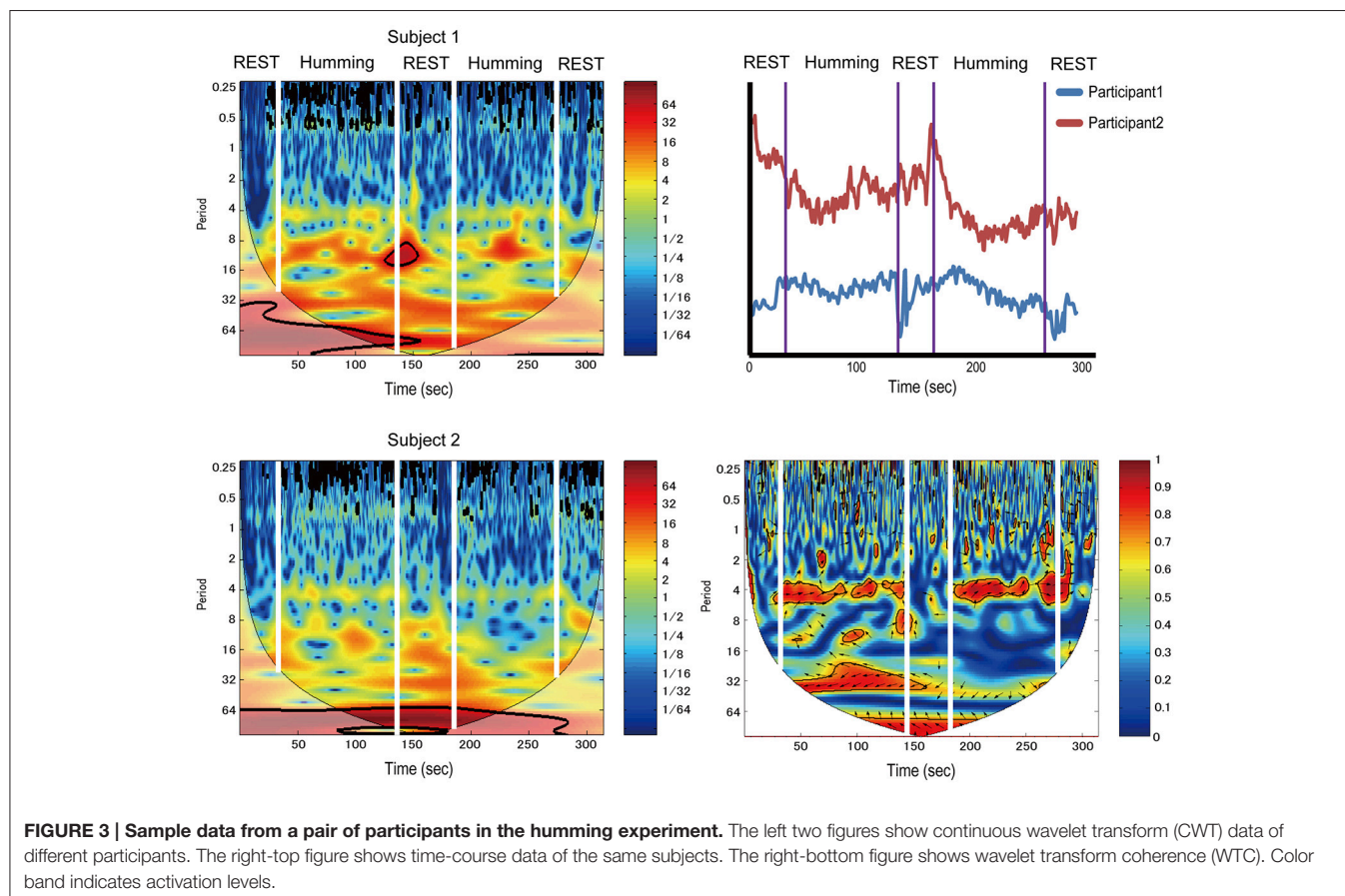
In the humming experiment, the procedure was identical to the singing experiment except that participants hummed the songs.

As the present experiment employed a three (participant 1 singing/humming, participant 2 singing/humming, and both

participants singing/humming)  $\times$  3 (first FtF, FtW, and second FtF) within-subject design, a total of nine values for coherence increase was obtained in each pair.

### fNIRS Data Acquisition

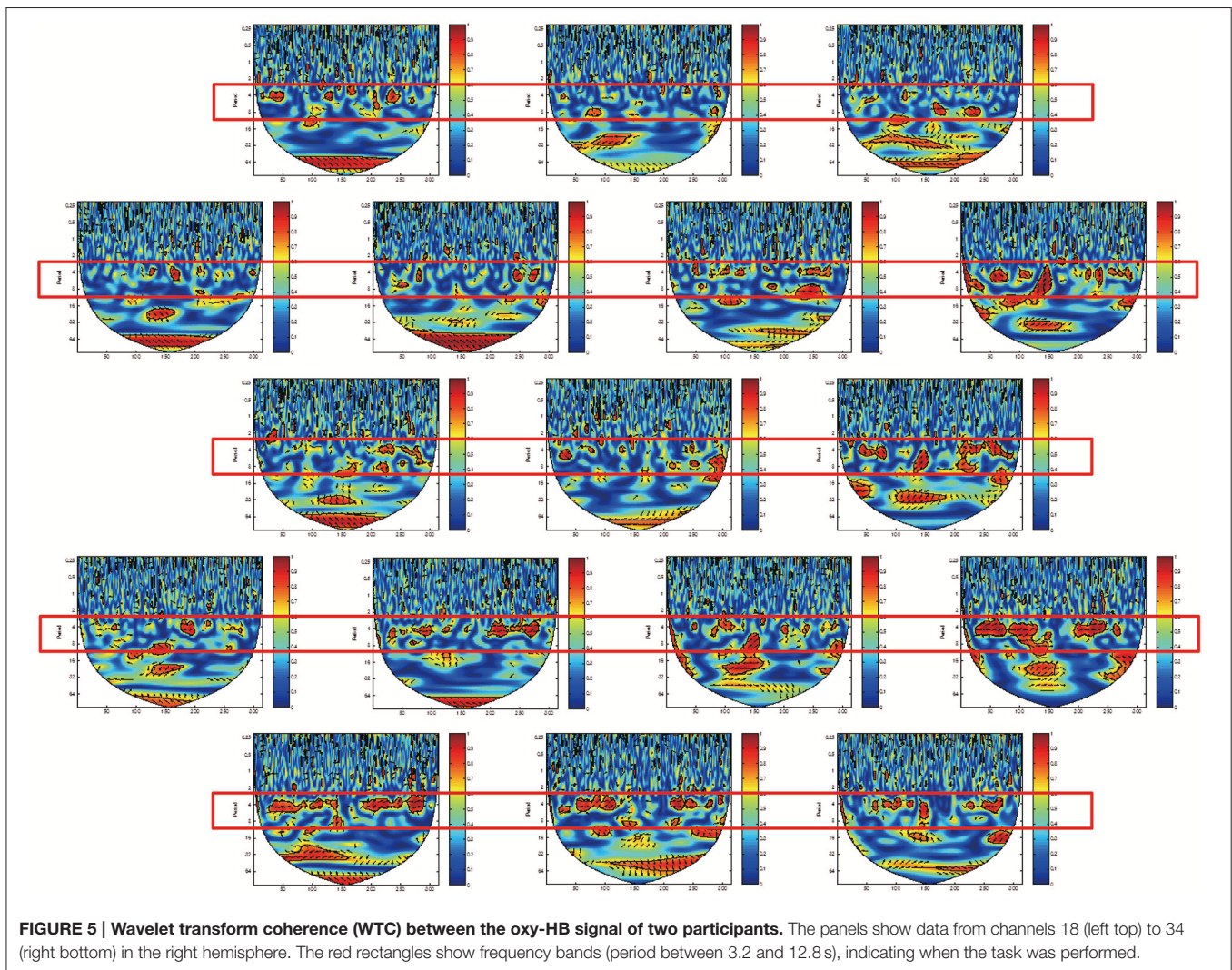
For the NIRS data acquisition, we employed a multichannel high-speed LABNIRS (Shimadzu, Japan) near-infrared spectroscopy measuring system to measure concentration variations in oxy-Hb and deoxy-Hb with easy operation. Combination of a three wavelength emitting system (780, 805, and 830 nm infra-red peak wavelengths) and a coupled photomultiplier detector tube achieved excellent sensitivity with scalability to increase the



caused mainly by oxy-Hb and deoxy-Hb, which have different absorption spectra, and the isosbestic point is in the vicinity of 805 nm. Therefore, if the molecular absorption coefficients of oxy-Hb and deoxy-Hb are known, the change in oxy-Hb and deoxy-Hb concentrations can be calculated by measuring the variation in absorption at two or more wavelengths. LABNIRS provides higher spatial resolution for high-density measurements and captures rapid cerebral blood flow signals in just 6 ms as compared with conventional sampling methods (<http://www.shimadzu.com/an/lifescience/imaging/nirs/nirs3.html>). A single  $3 \times 4$  cm measurement patch was attached to a whole-head fiber holder (Flexible Adjustable Surface Holder; FLASH), which was placed on each participant's head so that the fronto-temporal cortex and neighboring parietal cortex activity could be measured (Figure 2). We selected L-shaped fibers for the measurement, and the patch was positioned symmetrically over each participant's right and left brain (Figure 2). For example, red 5 and blue 5 in the left brain (channel 15) indicate emitter and detector, respectively (Figure 2C). Bottom channels (15–17 in the left brain and 34–32 in the right brain) were aligned to the Ca-Cp line (Talairach and Tournoux, 1988). Thus, in each patch, 12 emitters and 12 detectors were placed in the left and right brain of each participant, respectively, so that a total of 24 probes resulting in 34 measurement channels was employed for each participant. The sampling frequency we employed was 50 Hz.

number of channels (up to 142 channels) according to the purpose and number of participants connected to a single machine. The absorption in these wavelength regions are





## Data Analysis

To analyze synchronization in the fNIRS data, we employed wavelet transform coherence (WTC) analysis to evaluate the relationships between the fNIRS signals generated by a pair of participants by calculating the cross-correlation as a function of frequency and time (Torrence and Compo, 1998). WTC shows the local correlation between two time series data (Cui et al., 2012). We used the wavelet coherence package (Grinsted et al., 2004) provided at the website: <http://www.pol.ac.uk/home/research/waveletcoherence/>. WTCs were performed in each channel across two participants, focusing on oxy-HB signals in accordance with Cui et al. (2012). For the analysis, we re-sampled oxy-HB time-series data to 10 Hz in each channel, simply averaging five consecutive data points.

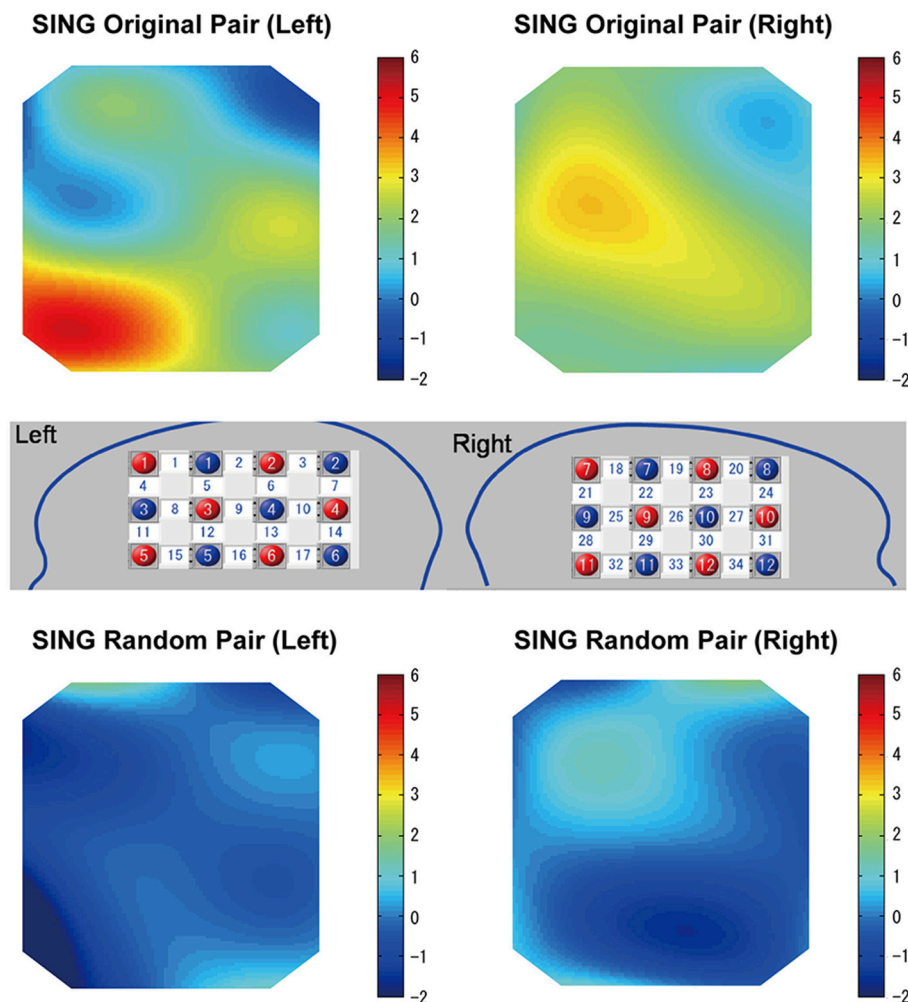
## Random Pair Analysis

To exclude the possibility that the obtained coherence increase in cooperative singing/humming relative to single singing/humming was due to the two participants being engaged in the same task in the cooperative conditions but not in

the single conditions, we performed a random pair analysis. The procedure was similar to that in Jiang et al. (2012), who tested coherence increase while two individuals were engaged in verbal communication. We selected two individuals from different pairs but sang the same song. Fifteen random pairs were made for the singing experiment and fourteen for the humming experiment. As the task duration differed across pairs, we adjusted the time-course data to be equal across the two individuals. That is, we specified the onset of singing/humming in each participant and defined the 30-s data before the onset as the pre-rest period and the 100-s data after the onset as the task period. We also specified the offset of singing/humming and defined the 30-s period after the offset as the middle- or post-rest period. WTC was applied to the two individual time-course data, and coherence increase was computed using the procedure described above.

For the random pair analysis, we determined the onset and offset of singing/humming and the rest period based on predetermined cue signals in the record. Therefore, the timing of singing/humming was matched between random pairs.





**FIGURE 6 | Heat maps of WTC group data comparing cooperative and single singing for the subject pairs singing together (top).** Heat maps of WTC group data comparing cooperative and single singing for randomly generated subject pairs (bottom). Numbers in the white squares indicate the channels between the emitters and detectors.

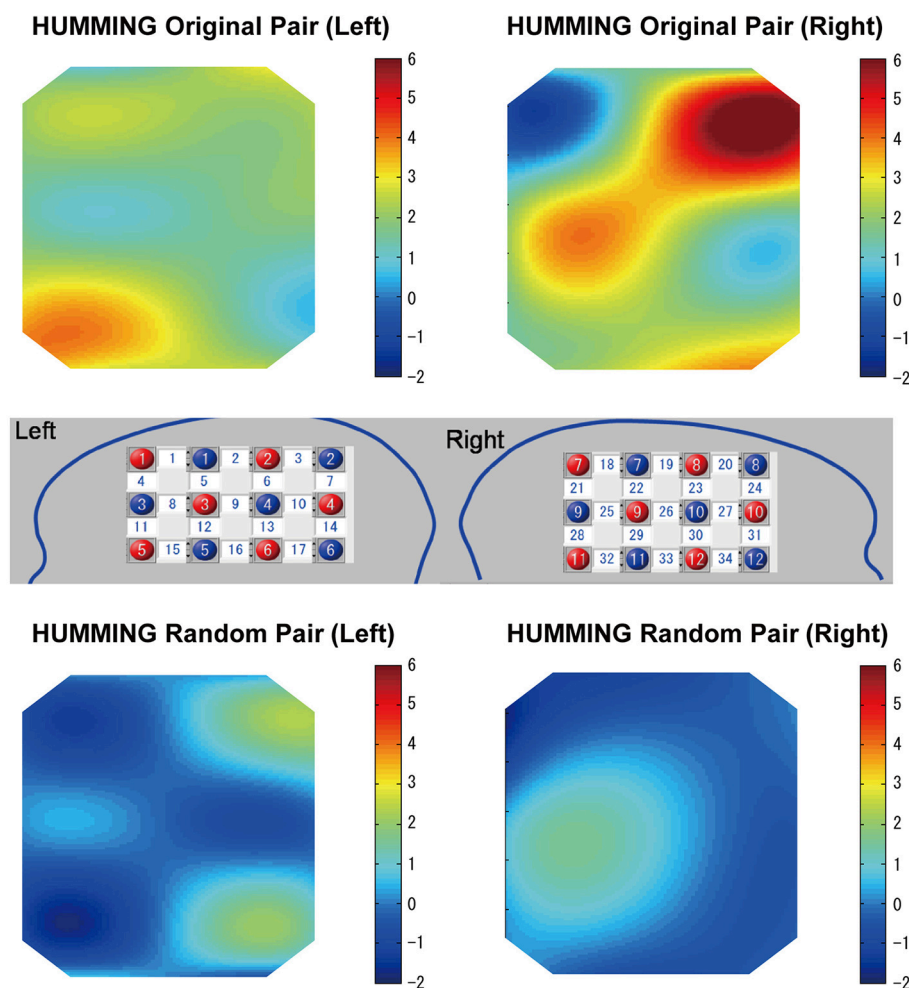
## RESULTS

Data from a pair of participants in the humming experiment are shown in **Figure 3**. The left two figures show continuous wavelet transform (CWT) data of different participants. The right-top figure shows the time-course data from both participants. The right-bottom figure shows WTC. WTC between participants is meaningful if the CWT of each participant does not show change between the rest and task intervals, although our data shown in **Figure 3** (left two figures) tended change a little at 4 s because the respiration changes CWT at 4 s.

We identified a frequency band that indicates the task was performed at approximately between 3.2 and 12.8 s (corresponding to a frequency of 0.3–0.08 Hz; **Figures 4, 5**, red rectangles). Cui et al. (2012) found a similar frequency band from data using a cooperative task in which the difference between the response times of both participants was smaller than a threshold time. Their frequency band includes the period of the

trial (7 s), indicating that the coherence increase in their band is task-related. We assumed our cooperated singing/humming conditions were similar to their cooperative game. In our study, breathing of both participants played a critical role in synchronized singing, and singing occurs only when the breathing occurred at a specific period of about 4 s (frequency of respiration at about 15 breaths per min; Vaschillo et al., 2006). Right-, left-, and downward arrows indicates in-phase, out-of-phase, and direction of WTC between the raw oxy-HB signal of two participants, respectively (**Figures 4, 5**).

Coherence across two participants in the task phase (e.g., cooperative singing) was computed by averaging the coherence values of two singing blocks where two participants sang together for about 100 s. As Cui et al. (2012) suggested, we defined a coherence increase as the averaged coherence value in two task blocks minus the average coherence value in the rest block. That is, the averaged coherence value in the rest condition was subtracted from that of the singing/humming conditions, and the



**FIGURE 7 | Heat maps of cooperative (original) humming and single humming (group data).** In the cooperative pair, stronger coherence was obtained in the bilateral inferior frontal cortices, the right middle frontal cortex and middle temporal cortex (**top**). This result was not observed for the random pair (**bottom**). Red and blue circles indicate the positions of the emitters and detectors in the left and right head, respectively. Numbers in the white squares indicate the channels between the emitters and detectors.

difference was used as an index of the neural synchronization increase between partners. Coherence increases were analyzed with a repeated ANOVA, including two factors with three levels in each. Because we repeated  $F$ -tests over 34 channels,  $p$ -values for the main effects and interactions were adjusted using the FDR method ( $p < 0.05$ ).

### Coherence Under the Singing

In the cooperative pair, the coherence increase was greater in the cooperative singing condition in the left IFC and the right middle temporal cortex (**Table 1**) than in the single singing condition regardless of FtF or FtW. Inclusion of the opaque partition did not weaken the coherence in the cooperative singing. A repeated ANOVA showed a main effect of the type of singing in the left IFC (Ch 11 and 12) and the right middle temporal cortex (Ch 25) ( $P < 0.05$ , FDR corrected for multiple comparisons). The main effect was attributed to the greater coherence increase in the cooperative condition, as *post-hoc* multiple comparisons showed

a significant increase in the cooperative condition compared with the single condition.

**Figure 6** (top) shows heat maps in which the coherence increase in the cooperative singing condition was compared with the single singing condition using a one-sample  $t$ -test for each channel. For the maps, we averaged the coherence increase in the cooperative condition and in the single condition (subject 1-sing and subject 2-sing) across the three visibility conditions (first FtF, FtW, and second FtF). Therefore, the heat maps correspond to T-maps smoothed by a spline correction method, which illustrates the channels that showed greater coherence increase in the cooperative condition than in the single condition. However, the coherence increase was equivalent across visibility (FtF vs. FtW) conditions (**Figure 8**). An ANOVA of data from the left IFC did not show a main effect of the visibility,  $F_{(2, 28)} = 0.80$ ,  $p = 0.46$ ,  $\eta_p^2 = 0.05$ . Similarly of the right IFC, a main effect of visibility was not significant,  $F_{(2, 28)} = 1.46$ ,  $p = 0.25$ ,  $\eta_p^2 = 0.09$ , either.

**TABLE 1 | Brain regions activated under the cooperative condition as compared with the single condition.**

Channel	Region	R/L	F-values	$\eta^2_p$
<b>HUMMING EXPERIMENT</b>				
3	Parietal cortex	L	6.50	0.34
11	Inferior frontal cortex	L	5.24	0.29
12	Inferior frontal cortex	L	7.87	0.37
15	Inferior frontal cortex	L	11.02	0.46
17	Inferior temporal cortex	L	6.20	0.32
23	Middle frontal cortex	R	15.56	0.55
24	Middle frontal cortex	R	5.80	0.31
25	Middle temporal cortex	R	6.81	0.34
29	Middle temporal cortex	R	5.03	0.28
31	Inferior frontal cortex	R	6.10	0.32
32	Inferior temporal cortex	R	5.46	0.30
33	Inferior temporal cortex	R	7.13	0.35
34	Inferior frontal cortex	R	4.85	0.27
<b>SINGING EXPERIMENT</b>				
11	Inferior frontal cortex	L	11.55	0.45
12	Inferior frontal cortex	L	7.45	0.35
25	Middle temporal cortex	R	6.64	0.33

In the random pair, a coherence increase in the cooperative condition was not found (**Figure 5**, bottom). A repeated ANOVA did not show a main effect of singing type in all channels, nor a main effect of the visibility ( $P > 0.05$ ). The interaction between factors was not significant in all channels ( $P > 0.05$ ).

## Coherence Under the Humming

Similar to the singing experiment, in the cooperative pair, the coherence increase was greater in the cooperative humming condition than in the single humming condition, but in more brain areas, including the left parietal cortex, the bilateral IFC, the right middle frontal cortex, and the right middle temporal cortices (**Table 1**). **Figure 4** shows WTC between the oxy-HB signal of two participants from channel 32 in the right hemispheres. **Figure 5** shows the other WTC data from channels 18–34 in the right hemispheres. The coherence increase was independent of the visibility condition. A repeated ANOVA showed main effects of the type of humming in the bilateral IFC (Ch11, 12, 15, 34), the left middle frontal cortex (Ch23, 24), the right parietal cortex (Ch3), the right middle temporal cortex (Ch 25, 29) and the bilateral inferior temporal cortex (Ch17, 32, 33;  $P < 0.05$ , FDR corrected for multiple comparisons). Like the singing experiment, we made heat maps (**Figure 6**) that compared the coherence increase in the cooperative condition and in the single condition across the visibility conditions, finding a stronger coherence increase in the cooperative humming condition relative to the single condition. A main effect of the visibility condition was not significant in all channels ( $P > 0.05$ ), and no interaction was obtained in all channels ( $P > 0.05$ ).

In the humming experiment, however, the coherence increase in the left IFC was greater in the second FtF than in the FtW

condition, but not in the right IFC (**Figure 8**). An ANOVA of the left IFC showed a significant main effect of the visibility,  $F_{(2, 26)} = 5.12$ ,  $p = 0.01$ ,  $\eta^2_p = 0.28$ , and Sheffer's modified sequentially Bonferroni test showed a significant difference in coherence between the second FtF condition and FtW condition ( $p = 0.02$ ). Such a main effect was not found in the right IFC,  $F_{(2, 26)} = 0.34$ ,  $p = 0.73$ ,  $\eta^2_p = 0.02$ .

In the random pair, the coherence increase in the cooperative condition did not differ from that in the single condition (**Figure 7**). The repeated ANOVA neither showed main effects of humming type in all channels nor a main effect of the visibility conditions ( $P > 0.05$ ). The interaction between factors was not significant in all channels ( $P > 0.05$ ).

## Coherence Under FtF and FtW

Because Jiang et al. (2012) found the left IFC (Ch 15) had stronger coherence in FtF, we examined this region and compared the coherence increase in cooperative singing/humming among the visibility conditions. In addition, the right IFC, which was in the equivalent position (Ch 34), was also analyzed. The coherence increase in the cooperative condition was compared among the visibility conditions. One-way repeated ANOVA was performed, and Sheffer's modified sequentially Bonferroni test was applied for multiple comparisons when a significant main effect was detected.

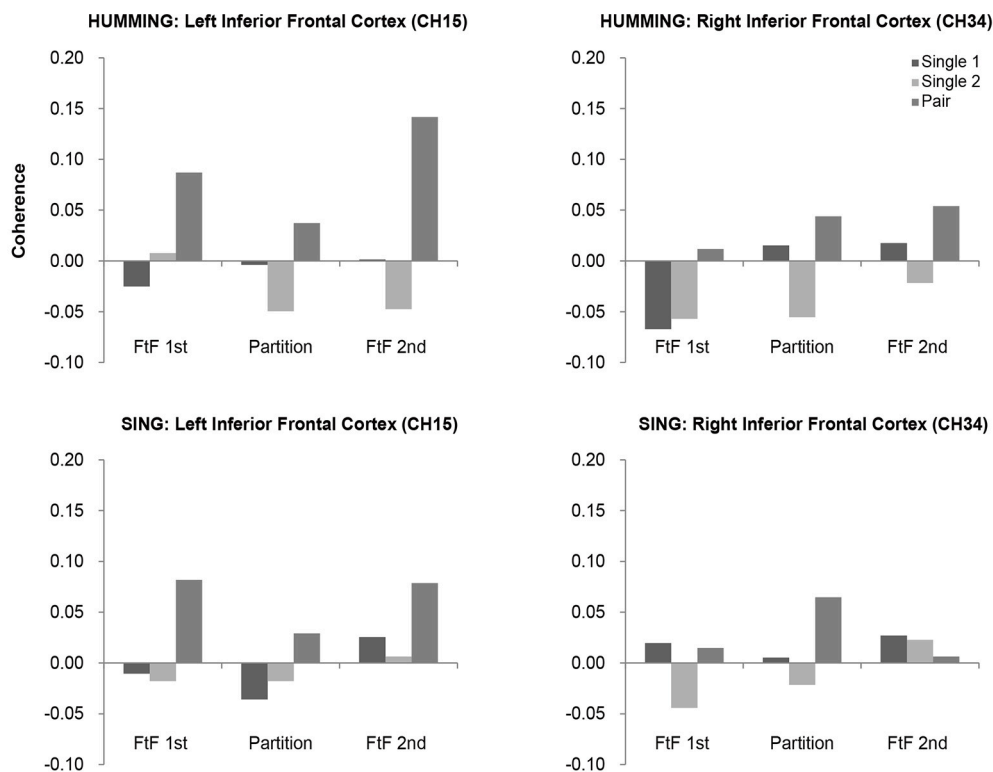
## Control Experiment

In order to check whether our results were influenced by the task, we conducted an earphone control experiment in which the humming was presented through earphones (the sound was attenuated so that only the person with the earphones could hear the humming;  $n = 28$ ; 14 pairs in the FtF condition in two sessions). This design could exclude the possibility that the synchronized activity was task-related. An ANOVA of this earphone experiment showed no main effect of the type of humming across all channels after applying false-discovery rate (FDR) correction ( $p < 0.05$ ).

## DISCUSSION

### Right and Left IFC

In the present study, we employed a single fNIRS machine to measure the neural synchronization of two participants simultaneously while they were singing or humming together. We found, by applying inter-brain WTC analysis between two interacting brains, a significant increase in the neural synchronization in the left IFC. Interestingly, the left IFC is where Broca's area is located, which has been identified as key to singing (Brodmann Area BA44 and 45 of the dominant brain). Broca's area contributes to the utterance of the words of a song. Similar activation in Broca's region in the left IFC has been observed during face-to-face dialog between partners (Jiang et al., 2012). Along with the left IFC, the right IFC was activated during humming, which was attributed to a coordinated production of melody. Cui et al. (2012) found significant neural synchronization in the left IFC



**FIGURE 8 | Coherence values in the bilateral inferior frontal cortices (IFC) in the humming and singing experiments.** In the humming experiment, the bilateral IFC showed greater coherence in the cooperative condition than in the single condition across sessions (FtF and FtW). In the singing experiment, greater coherence in the cooperative condition was found only in the left IFC.

during cooperation but not during competition. The current study confirmed no WTC during single or random-paired conditions, which clearly indicates bilateral IFC activation in cooperative tasks. Therefore, both the left and right IFC are likely responsible for synchronizing two brains, with activation of the left IFC being superimposed for the bias of verbal expression. Along with the right IFC, the middle temporal cortex and middle frontal cortex are suggested to contribute to neural synchronization during humming. However, the activation of the superior frontal cortex reported by Cui et al. (2012) under the cooperation task was not observed, likely due to the difference in tasks.

Recent studies using hyperscanning to investigate the temporal and emotional aspects of music production have been reported (Lindenberger et al., 2009; Babiloni et al., 2011; Babiloni and Astolfi, 2014). Using EEG-based hyperscanning, Lindenberger et al. (2009) reported increased brain activity in the theta frequency band (4–7 Hz) of the prefrontal cortex during synchronous music production with the help of a metronome. Similarly Babiloni et al. (2012) simultaneously recorded the brain activities of saxophonists playing music in an ensemble and reported a correlation between empathy and alpha desynchronization in the right ventral-lateral frontal gyrus (BA 44/45). Our findings of activation in the right IFC under cooperative humming are in good agreement with

these data. However, we found inter-brain oscillatory frequency bands of 0.3–0.08 Hz. The difference in frequencies can be explained in terms of the slow hemodynamic delay of about 3 s measured by fNIRS as compared with the fast waves measured by EEG.

## Face-to-face Cooperation

FtF social interactions are likely critical for synchronizing cooperation. Our study revealed FtF relatively tended to enhance activity of the left IFC under humming (Figure 8). FtW, however, showed negligible effects on the IFC and neighboring brain regions. These results are in good agreement with Jiang et al. (2012), who reported a significant increase in neural synchronization in the left IFC under FtF dialog, but not during back-to-back dialog, FtF monolog, or back-to-back monolog. In addition, we found that FtF played a critical role under humming, while FtW had negligible influence partly due to the importance of vocal rather than facial cooperation. As for why an increase in synchronization of the right IFC was seen for cooperative humming, only it could be that singing created a cognitive load.

A related study by Saito et al. (2010) that used fMRI reported pair-specific correlations of intrinsic brain activity during facial (eye) contact compared with non-paired subjects who were not in eye contact. They used an experimental paradigm in which



the participants could recognize the gaze of the other on a screen on which there was also depicted other objects. Their results suggested that the right IFC was active in couples during conditions like FtF in our study.

## Social Perspective

Cooperative singing may be beneficial to people whose sense of shared cooperation is weak. By singing together, an out-of-tune individual could be harmonized with an in-tune other, thus sharing joy through synchronized cooperation. Shared cooperation indicates the ability to create with others joint interactions and synchronized attention underlaid by cooperative motives (Tomasello, 2009). Furthermore, singing together enhances emotional relief and pleasure, and is expected to yield a sense of mutual trust and cooperation (Gaston, 1968; Anshel and Kipper, 1988).

Cooperative singing could also be partly interpreted as the result of mutual activations in the human mirror neuron system (MNS) of the prefrontal regions of two people. People have a tendency to imitate others using the MNS in order to conform to an indicator of group identity. Moreover, the MNS is likely located in the IFC and adjacent ventral premotor areas (Rizzolatti and Arbib, 1998; Iacoboni and Dapretto, 2006).

## REFERENCES

- Anshel, A., and Kipper, D. A. (1988). The influence of group singing on trust and cooperation. *J. Music Therapy* 25, 145–155. doi: 10.1093/jmt/25.3.145
- Astolfi, L., Toppi, J., De Vico Fallani, F., Vecchiato, G., Cincotti, F., Wilke, C. T., et al. (2011). Imaging the social brain by simultaneous hyperscanning during subject interaction. *IEEE Intell. Syst.* 26, 38–45. doi: 10.1109/MIS.2011.61
- Babiloni, C., Buffo, P., Vecchio, F., Marzano, N., Del Percio, C., Spada, D., et al. (2012). Brains on concert: frontal oscillatory alpha rhythms and empathy in professional musicians. *Neuroimage* 60, 105–116. doi: 10.1016/j.neuroimage.2011.12.008
- Babiloni, C., Vecchio, F., Infarinato, F., Buffo, P., Marzano, N., Spada, D., et al. (2011). Simultaneous recording of electroencephalographic data in musicians playing in ensemble. *Cortex* 47, 1082–1090. doi: 10.1016/j.cortex.2011.05.006
- Babiloni, F., and Astolfi, L. (2014). Social neuroscience and hyperscanning techniques: past, present and future. *Neurosci. Biobehav. Rev.* 44, 76–93. doi: 10.1016/j.neubiorev.2012.07.006
- Babiloni, F., Cincotti, F., Mattia, D., Mattiocco, M., De Vico Fallani, F., Tocci, A., et al. (2006). Hypermethods for EEG hyperscanning. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 1, 3666–3669. doi: 10.1109/IEMBS.2006.260754
- Bailey, W. J. (2003). Insect duets: underlying mechanisms and their evolution. *Physiol. Entomol.* 28, 157–174. doi: 10.1046/j.1365-3032.2003.00337.x
- Csikszentmihalyi, M. (1990). *Flow: The Psychology of Optimal Experience*. New York, NY: Harper.
- Cui, X., Bryant, D. M., and Reiss, A. L. (2012). NIRS-based hyperscanning reveals increased interpersonal coherence in superior frontal cortex during cooperation. *Neuroimage* 59, 2430–2437. doi: 10.1016/j.neuroimage.2011.09.003
- De Vico Fallani, F., Nicosia, V., Sinatra, R., Astolfi, L., Cincotti, F., Mattia, D., et al. (2010). Defecting or not defecting: how to “read” human behavior during cooperative games by EEG measurements. *PLoS ONE* 5:e14187. doi: 10.1371/journal.pone.0014187
- Dunbar, R. (2010). *How Many Friends Does One Person Need?: Dunbar's Number and Other Evolutionary Quirks*. Boston, MA: Harvard University Press.
- Fortune, E. S., Rodríguez, C., Li, D., Ball, G. F., and Coleman, M. J. (2011). Neural mechanisms for the coordination of duet singing in wrens. *Science* 334, 666–670. doi: 10.1126/science.1209867
- Frith, C. D., and Frith, U. (1999). Interacting mind - a biological basis. *Science* 286, 1692–1695. doi: 10.1126/science.286.5445.1692
- Gaston, E. T. (1968). *Music in Therapy*. New York, NY: Macmillan.
- Geissmann, T. (2002). Duet-splitting and the evolution of gibbon songs. *Biol. Rev.* 77, 52–76. doi: 10.1017/S1464793101005826
- Grinsted, A., Moore, J. C., and Jevrejeva, S. (2004). Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlin. Processes Geophys.* 11, 561–566. doi: 10.5194/npg-11-561-2004
- Gunji, A., Ishii, R., Chau, W., Kakigi, R., and Pantev, C. (2007). Rhythmic brain activities related to singing in humans. *Neuroimage* 34, 426–434. doi: 10.1016/j.neuroimage.2006.07.018
- Iacoboni, M., and Dapretto, M. (2006). The mirror neuron system and the consequences of its dysfunction. *Nat. Rev. Neurosci.* 7, 942–951. doi: 10.1038/nrn2024
- Jiang, J., Dai, B., Peng, D., Zhu, C., Liu, L., and Lu, C. (2012). Neural synchronization during face-to-face communication. *J. Neurosci.* 32, 16064–16069. doi: 10.1523/JNEUROSCI.2926-12.2012
- Kimura, D. (1964). Left-right difference in the perception of melodies. *Q. J. Exp. Psychol.* 16, 355–358. doi: 10.1080/17470216408416391
- King-Casas, B., Tomlin, D., Anen, C., Camerer, C. F., Quartz, S. R., and Montague, P. R. (2005). Getting to know you: reputation and trust in a two-person economic exchange. *Science* 308, 78–83. doi: 10.1126/science.1108062
- Lindenberger, U., Li, S., Gruber, W., and Müller, V. (2009). Brains swinging in concert: cortical phase synchronization while playing guitar. *BMC Neurosci.* 10:22. doi: 10.1186/1471-2202-10-22
- Mithen, S. (2005). *The Singing Neanderthals: The Origins of Music, Language and Body*. London: Weidenfeld & Nicholson.
- Montague, P. R., Berns, G., Cohen, J. D., McClure, S. M., Pagnoni, G., Dhamala, M., et al. (2002). Hyperscanning: simultaneous fMRI during linked social interactions. *Neuroimage* 16, 1159–1164. doi: 10.1006/nimg.2002.1150
- Rizzolatti, G., and Arbib, M. A. (1998). Language within our grasp. *Trends Neurosci.* 21, 188–194. doi: 10.1016/S0166-2236(98)01260-0

It is not surprising then that cooperative singing, which is a form of collective experience, gives rise to neural synchronization.

In summary, we examined how two brains make one synchronized mind using cooperative singing/humming between two people and hyperscanning. Hyperscanning allowed the observation of dynamic cooperation in which participants interacted with each other. We used fNIRS to record the brain activity of two brains while they cooperatively sang or hummed a song in FtF or FtW conditions. Inter-brain WTC between the two interacting brains showed a significant increase in the neural synchronization of the left IFC for both singing and humming regardless of FtF or FtW compared with singing or humming alone. On the other hand, the right IFC showed an increase in neural synchronization for humming only. Our data suggest, the application of hyperscanning during cooperative tasks could improve understanding of social cooperation.

## ACKNOWLEDGMENTS

The study was supported in part by the Grants #22220003, #15H01690 to NO and #23240036 to MO from JSPS (Japan Society for the Promotion of Science).

- Saito, D. N., Tanabe, H. C., Izuma, K., Hayashi, M. J., Morito, Y., Komeda, H., et al. (2010). Stay tuned: inter-individual neural synchronization during mutual gaze and joint attention. *Front. Integr. Neurosci.* 5:127. doi: 10.3389/fnint.2010.00127
- Talairach, J., and Tournoux, P. (1988). *Co-planar Stereotaxic Atlas of the Human Brain*. New York, NY: Thieme.
- Tomasello, M. (2009). *Why We Cooperate*. Cambridge: MIT Press.
- Torrence, C., and Compo, G. P. (1998). A practical guide to wavelet analysis. *Bull. Am. Meteorol. Soc.* 79, 61–78.
- Vaschillo, E. G., Vaschillo, B., and Lehrer, P. M. (2006). Characteristics of resonance in heart rate variability stimulated by biofeedback. *Appl. Psychophysiol. Biofeedback* 31, 129–142. doi: 10.1007/s10484-006-9009-3

**Conflict of Interest Statement:** 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.

Copyright © 2015 Osaka, Minamoto, Yaoi, Azuma, Minamoto Shimada and Osaka. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# What Happens in a Moment

Mark A. Elliott<sup>1\*</sup> and Anne Giersch<sup>2</sup>

<sup>1</sup> School of Psychology, National University of Ireland Galway, Galway, Ireland, <sup>2</sup> INSERM U1114, Department of Psychiatry, Fédération de Médecine Translationnelle de Strasbourg, Strasbourg University Hospital, Strasbourg, France

There has been evidence for the very brief, temporal quantization of perceptual experience at regular intervals below 100 ms for several decades. We briefly describe how earlier studies led to the concept of “psychological moment” of between 50 and 60 ms duration. According to historical theories, within the psychological moment all events would be processed as co-temporal. More recently, a link with physiological mechanisms has been proposed, according to which the 50–60 ms psychological moment would be defined by the upper limit required by neural mechanisms to synchronize and thereby represent a snapshot of current perceptual event structure. However, our own experimental developments also identify a more fine-scaled, serialized process structure within the psychological moment. Our data suggests that not all events are processed as co-temporal within the psychological moment and instead, some are processed successively. This evidence questions the analog relationship between synchronized process and simultaneous experience and opens debate on the ontology and function of “moments” in psychological experience.

## OPEN ACCESS

### Edited by:

Lihan Chen,  
Peking University, China

### Reviewed by:

Michael Herzog,  
Swiss Federal Institute of Technology  
in Lausanne, Switzerland

Hansem Sohn,  
Massachusetts Institute of  
Technology, USA

### \*Correspondence:

Mark A. Elliott  
mark.elliott@nuigalway.ie

### Specialty section:

This article was submitted to  
Perception Science,  
a section of the journal  
Frontiers in Psychology

**Received:** 12 July 2015

**Accepted:** 25 November 2015

**Published:** 07 January 2016

### Citation:

Elliott MA and Giersch A (2016)  
What Happens in a Moment.  
Front. Psychol. 6:1905.  
doi: 10.3389/fpsyg.2015.01905

**Keywords:** time, psychological moment, perceptual organization, serial processing, Simon effect

## WHAT HAPPENS IN A MOMENT

On an experiential level, a perceptual moment is usually defined as what one experiences in the immediate and “specious present,” i.e., a time interval spanning several hundreds of milliseconds (ms) to a second (see Anderson and Grush, 2009, for definition). For example, when listening to a melody, it would correspond to what is presently in mind, including the note played just before and possibly the note expected to immediately follow. Events are thus clearly distinguished in time within the experienced present. However, a shorter interval has been recorded related to the discretization of psychological events. In this paper we will describe the evidence for this shorter interval. We first describe how earlier work led to the idea that events are processed as co-temporal within elementary time windows of 50–60 ms. We then review our data and the literature that challenges this view by showing that events are automatically distinguished in time at shorter asynchronies. This will allow us to discuss the structure that brings about an elementary quantization of perceptual events.

Although not the first to postulate its existence, Brecher was amongst the first to empirically define a psychological moment below 100 ms. In an ingenious set of experiments, Brecher (1932) established the minimal time required for the perceptual separability of two or more events presented repeatedly and in sequence. Importantly, Brecher’s estimate was near identical across modalities in healthy adult participants at 55.3 ms for tactile stimulation and 56.9 ms for visual stimulation, with standard deviations of no greater than 1.4 ms across 14 subjects. Also important was the close corroboration of Brecher’s empirical estimate with earlier, although difficult to verify, estimates given by Lalanne (1876) and von Baer. von Baer (1864) is believed to have proposed to the Russian Academy of Sciences in St. Petersburg a fundamental quantum of experienced time at

$1/18$ th of a second, deviations from which could allow for the accurate prediction of the life span of the organism. While influential in the development of ideas such as that of “Umwelt” (phenomenal surrounding) proposed by von Uexküll (1957), no reference is made to this topic in proceedings of the St. Petersburg meeting (e.g., Von Baer, 1909).

A more contemporary conceptualization is related to the idea that during this 53–55 ms interval, neural mechanisms are engaged that render two stimuli as the parts of a single event structure. In this context, the interval described here may refer to a minimum number of oscillations (and so maximum interval) required for two or more neurons to form an assembly that allows for the coding of perceptual structure. This idea refers to a theory expressed in the context of literature on oscillatory neural binding (i.e., the neural code believed to be responsible for the neural coding of relations in perceptual structure, see Singer, 1993, 1999). This theory postulates a minimum number of oscillatory cycles between synchronized neurons is required for a synchronized assembly to be statistically separable from a spurious synchronization—and thus treated by the perceptual system as likely to be coding multi-dimensional perceptual structure. Given oscillations in the broad band 30–70 Hz have been shown to be associated with coding perceptual structure, this would entail between 2 and 3 synchronized events are sufficient to signal perceptual structure within a 53–55 ms interval. It must be acknowledged that this estimate should be treated as speculative, though, since estimates of binding oscillations are taken from a variety of species, which may be subject to different perceptual moments as compared to human beings (see Brecher, 1932, for examples). There are, however a number of estimates of simultaneity thresholds and before treating Brecher’s moment in greater detail these, and some other moments require consideration: minimum simultaneity thresholds have been estimated from reports of the simultaneity of spatially separate flashes or lines presented in close spatial proximity. Stimuli such as these may be perceived as simultaneous for inter-flash intervals within the range 1–5 ms; only at larger intervals do they yield the perception of successiveness (in this case of apparent motion, see Sweet, 1953; Westheimer and McKee, 1977; Wehrhahn and Rapf, 1992). Other estimates suggest maximum intervals for the perception of simultaneity and by extension minimum time differences in temporal order discrimination (with attendant motion perception) for intervals of between 17 and 44 ms (Exner, 1875). Empirical evidence has accumulated over the last decades showing that temporal order thresholds across modalities reliably lie in the time range between 20 and 60 ms (Pöppel, 1997; Fink et al., 2006; Babkoff and Fostick, 2013), and with some variation as a function of stimulus properties (Wittmann, 2011).

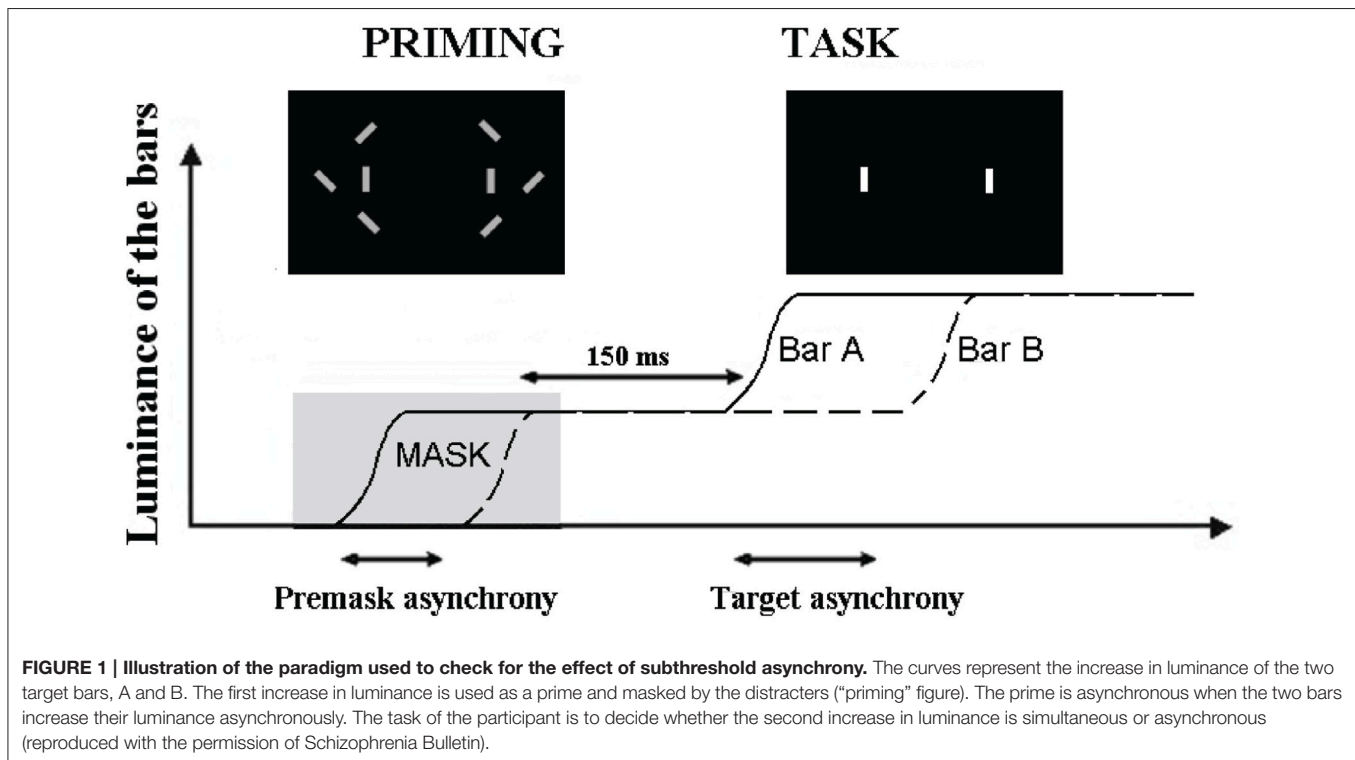
Unlike paradigms relying upon simultaneity judgments to two events presented simultaneously, or with a small asynchrony, Brecher presented his stimuli as a series of paired events and in this series of events, each event consisted of two simultaneous or slightly asynchronous stimuli. This design, while lending greater ecological validity to his estimate leads to the requirement to process temporal relationships—not only between the two stimuli in each event, but also across events

within the series. Simultaneity is estimable only if there is a third (or subsequent) event with which events within the simultaneity are perceptually separable (discussed in detail in Elliott et al., 2007). The point here is that “moments” are not only defined by stimuli bound into a simultaneity, but also by the segregation of the simultaneity from other events in past (and future) moments. This leads to a time paradox, i.e., our difficulty to understand how our perception can be both discontinuous, with separable events, and continuous, with a feeling that time flows without interruption with all events related in time. Considering our experience is rarely of events in staccato, the moments that are measured experimentally may not be a description of phenomenal experience itself but of an underlying discretization such as that implied by Brecher and proposed by Stroud (1955), who postulated the existence of 110 ms quanta underlying phenomenal experience. This was a reinterpretation of Allport’s (1964) study of perceived simultaneity in terms of the moment as a continuous, running sample of the input (a Traveling Moment Hypothesis) and thus reconciles the ideas of continuity and discontinuity. This is an important idea in the present context as it allows us to be clear that we are not discussing—directly—our experience of duration (aka time perception) or simultaneity or asynchrony. Instead, and as will become clear, we are discussing what we can learn of the discretization of event structure, implicitly (and very likely at a neural level although there exists very little direct data to support this), and how this relates to the experience of events in an uninterrupted, temporal continuity.

At the ceiling of estimates concerned with immediate perceptual experience are perhaps those of Efron (1970a,b), who describes the minimum duration of an experience as of 137 ms. However and because of its dependence on the organization of event structure, Brecher’s moment may be an estimate of an upper limit on elementary perceptual integration. This seems plausible by analogy to spatial organization. In this case, binding between separate features is generally held to result from neuronal assemblies formed by the synchronization of contributive neurons via phase alignment of their spiking in bursts of frequency oscillations (reviewed in Singer, 1999). In addition, Duncan and Humphreys (1989) showed that the very early spatial coding guiding activities such as visual search requires the preattentive segregation of target features from distractors, implying that early grouping is partly derived from relational coding across the entire visual scene.

Using a paradigm similar to that employed by Brecher, in that the paradigm employed repeating visual presentations, Elliott et al. (2007) found mean simultaneity thresholds to target pairings in very close proximity to those reported by Brecher (61 ms). One modification employed by Elliott et al. (2007) was the masked presentation of an asynchrony just prior to target presentation. This allowed investigation of whether a subthreshold synchrony ( $SB_S$ ) or asynchronies ( $SB_A$ ) would bring about a shift in the threshold for perceived simultaneity, and at which asynchronies, if any, this shift would be found. As illustrated in **Figure 1**, two stimuli were first presented synchronously ( $SB_S$ ) or asynchronously ( $SB_A$ ). The asynchrony was made non-detectable by embedding the





stimuli within distracters. These stimuli served as primes and, after the disappearance of distracters, they were increased in luminance, following which, participants had to decide whether this luminance increase was simultaneous or asynchronous. Interestingly,  $SB_S$  and  $SB_A$  produce different patterns of effects only for targets over a very short range of stimulus-onset asynchronies (SOAs) between targets (including physically simultaneous targets). For target SOAs of up to 21 ms, there appeared to be a small bias towards simultaneity judgments following exposure to  $SB_S$ , relative to simultaneity judgments following exposure to  $SB_A$ . In addition, and for presentations above threshold (61 ms) there seems to be a decreasing tendency to report simultaneity when the targets were preceded by  $SB_S$ . That enhanced simultaneity reportage (following  $SB_S$ ) is maintained for SOAs of 0–(14–21) ms is interesting in that the interval 14–21 ms is very close to the maximum separation in time between the firing of different neurons within synchronized neural assemblies in visual cortex (see, e.g., Gray et al., 1989, for data; and Singer, 1993, for review). The rhythmic synchronization of neuronal firing is believed to facilitate the formation of functional neuronal assemblies with those operating in the EEG gamma band (30–70 Hz) associated with functions that includes perceptual processing. What is suggested by the findings of Elliott et al. (2007) is that subthreshold stimulus asynchronies at very short SOAs may influence the efficiency of neuronal synchronization from which we can conclude that functional neuronal assemblies form within the moment with the goal of representing coherent perceptual structure.

The functional moment appears to be constrained by the temporal properties of neurons, i.e., the time needed to

synchronize neuron assemblies. In turn, it constrains perception by providing a temporal organization. Since moments are too short to yield a perception of duration (Wittmann, 2011), they are thought to be elementary elements in the composition of trains of thoughts. However, they may not correspond to elementary information: on the contrary and as implied above, they may form as a consequence of information integration. In case of multisensory information, up to 100–200 ms may be needed to distinguish an asynchrony between visual and auditory information (Vatakis and Spence, 2007; van Wassenhove, 2009). Does this mean that perception is a series of snapshots from which we rebuild an experiential continuity a posteriori (Neisser, 1967; Ullman, 1979; Shimojo, 2014; van Rullen et al., 2014)? This possibility requires us to understand both how visual information is correctly organized if it is initially integrated within elementary windows (Gepshtein and Kubovy, 2000), as well as how the coding of discrete moments can be reconciled with our experience of events as in continuous time. Several solutions have been proposed; for instance the overlap of moments (Dainton, 2010). However, there are many different conceptualizations (Phillips, 2014), and recent results may suggest alternative possibilities, which are discussed in the following.

Scharnowski et al. (2009) and Pilz et al. (2013) have found that stimuli perceived as fused in time (i.e., co-temporal) may in fact be initially processed as temporally segregated. In these studies, the authors used stimuli presented in sequence over short time intervals that lead to a temporally fused percept. They applied either TMS (Scharnowski et al., 2009) or masking (Pilz et al., 2013) at different delays to disturb information processing, and examined which of the two successive stimuli dominated

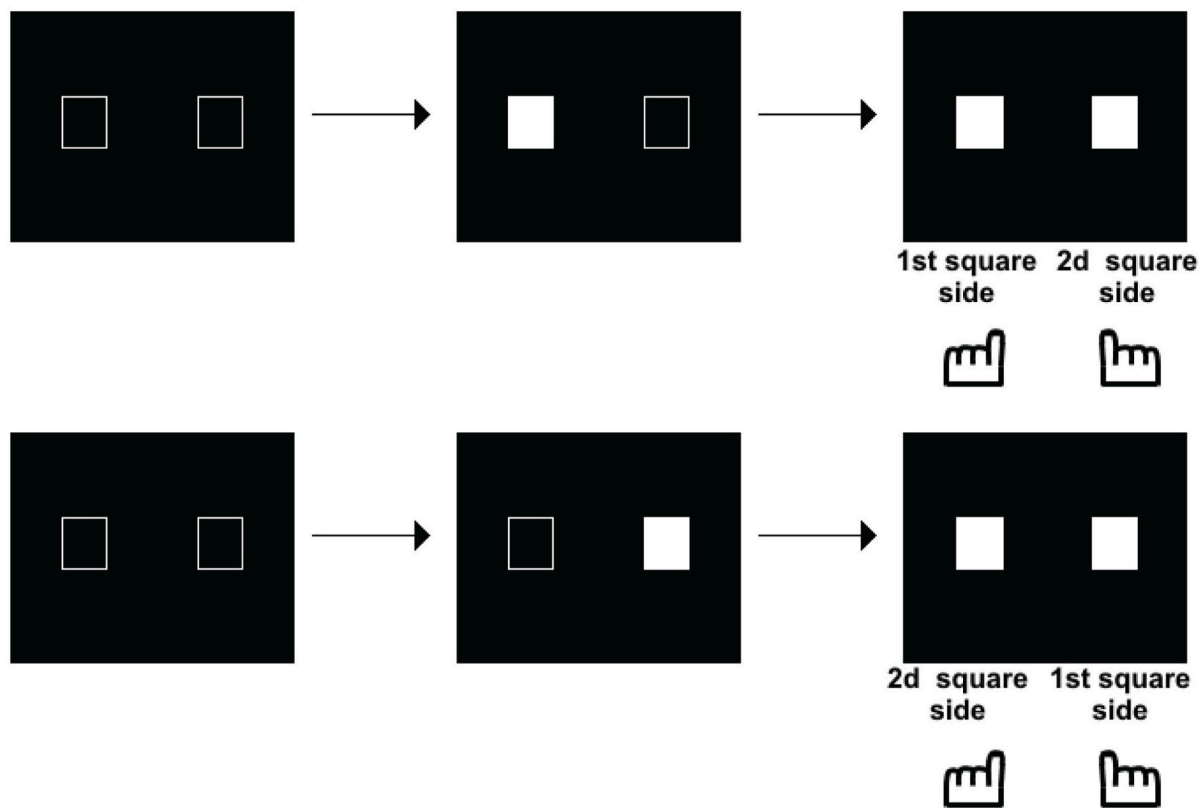
the perception. This procedure allowed them to establish that the processing of the two successive stimuli can be disturbed distinctly and in turn. Their results show that disturbance applied 45–90 ms after stimuli onset affects the processing of the first stimulus (leading to the dominance of the second stimulus), whereas disturbance applied 95–420 ms after stimuli onset affects the processing of the second stimulus (leading to the dominance of the first stimulus). It is only after delays of 400–500 ms that both stimuli are perceived as temporally fused with neither TMS nor masking modifying the fusion. These results show that information integration is slow and that perception is more discrete than believed from subjective reports. In addition, the results suggest a specific time course for information processing: for between 400 to 500 ms, successive stimuli are as yet not integrated, and instead processed one after the other, in sequence.

Such a possibility was explored in another series of studies, initially aimed at exploring the time course of perception in schizophrenia. These studies were motivated by the fact that patients with schizophrenia have been described as suffering from a fragmentation of consciousness, with a loss of the sense of time continuity (Fuchs, 2007). Several studies have shown a lengthening of the perceptual moment: i.e., patients required larger asynchronies than controls to detect asynchronies (Foucher et al., 2007; Giersch et al., 2009; Schmidt et al., 2011; Lalanne et al., 2012a; Martin et al., 2013). This effect was independent of decisional or another non-specific factor (reviewed in Giersch et al., 2013). The lengthening of the perceptual moment became quite large in presence of distracters (Giersch et al., 2009) or in case of multisensory signals (Martin et al., 2013). The integration of information within temporal windows of several 100 ms in patients questioned the way these subjects interact with the environment, especially as it contrasted with their mild pathological state. The implicit processing of stimuli over time was investigated, i.e., the ability to detect asynchronies independent of a conscious judgment. The Simon effect (Simon, 1969) was used to that aim, which corresponds to the tendency to press on the side of a stimulus independent of the task at hand. For example, if the task is to discriminate between squares and circles, and to press respectively on the left and right side in case of a square vs. a circle, subjects will tend to press on the left whenever the stimulus is displayed on the left, even if it is a circle. The mechanisms of this effect are reviewed in Hommel (2011a,b) and van der Lubbe and Abrahamse (2011), and it was used as a tool to examine the automatic processing of stimuli over time. This first required some adaptation of the Simon effect, since two stimuli and not only one, are presented during temporal tasks. As a matter of fact, when two stimuli are simultaneously displayed on the screen, one on the right side and the other on the left side of the screen, responses cannot be biased on either side, since information is perfectly symmetrical. A bias can be observed only in case of an asymmetry between right and left sides, which occurs in case of an asynchrony between the two stimuli. In case of a clear asynchrony, we have shown that subjects are biased to press on the response key located on the side of the second stimulus, whether it is on the left or on the right (Lalanne et al., 2012a,b; illustrated in **Figure 2**).

This shows that the Simon effect can be used with a simultaneity/asynchrony discrimination task. The critical analysis, however, regarded the exploration of the Simon effect in case of undetected asynchronies, for SOAs below 20 ms. Inasmuch as such asynchronies yield the same amount of “simultaneous” responses as perfect synchrony, they may have been expected to inhibit presentation of a Simon effect. If stimuli are processed as co-temporal, they would indeed yield symmetrical information on both sides on the screen, thus precluding any response bias to either side. This is not what was observed, however. In healthy subjects and for asynchronies as short as 17 ms, a bias to the side of the second stimulus was still observed. Importantly, in patients a bias was also observed to the side of the first stimulus (Lalanne et al., 2012a,b), and this was observed even for asynchronies as short as 8 ms (Giersch et al., 2015). Apart from the significance regarding schizophrenia pathophysiology (Martin et al., 2014), this result is important because it suggests a dissociation between the automatic processing of stimuli over time at delays below 20 ms, and the explicit ability to distinguish events in time. First, patients’ ability to detect asynchronies is disturbed, and there is thus a large gap between their ability to explicitly discriminate visual stimuli in time (threshold around 50 ms) and their implicit processing over time (8 ms). Second, the Simon effect is reversed at short (on the side of the first stimulus) and at large asynchronies (on the side of the second stimulus). The implicit ability to discriminate stimuli in time may play a special role in the processing of visual information, by providing the means to follow stimuli over time at an implicit level and with a high temporal accuracy. The possibility that stimuli are processed successively at an unconscious level is suggested by the studies of Scharnowski et al. (2009) and Pilz et al. (2013). Elliott et al. (2007) also suggested that the processing of short asynchronies can be modulated by prior temporal information, i.e., pairs of events whose simultaneity or asynchrony was made non-detectable by the presence of distracters. Asynchronies used to study the time course of information processing (Scharnowski et al., 2009; Pilz et al., 2013) were generally set at around 40 ms, similar to the asynchronies used for primers in Elliott et al. (2007). What the Simon effect at 17 ms brings in addition is evidence that information delayed by asynchronies below 30 ms is not treated as co-temporal by all processes. Specifically, processing, even at small delays is dependent upon stimulus order, suggesting a serialization of processing, even within very short processing windows. This idea is supported by a recent study using healthy volunteers carried out by Poncelet and Giersch (2015). These authors used a priming paradigm to investigate the impact of two (unmasked) primes delayed by 17 ms on the subsequent detection of a target, or on the ordering of two targets (**Figure 3**).

The aim of this experiment was to check if the primes facilitated or inhibited the detection of a target displayed in the location of the first or second prime<sup>1</sup>. Facilitation would suggest that attentional mechanisms has deployed to the prime location,

<sup>1</sup>Here the paradigm did not consist in the exploration of the influence of simultaneity/asynchrony on a temporal judgment. The paradigm was built to examine how the successive primes were processed in time.

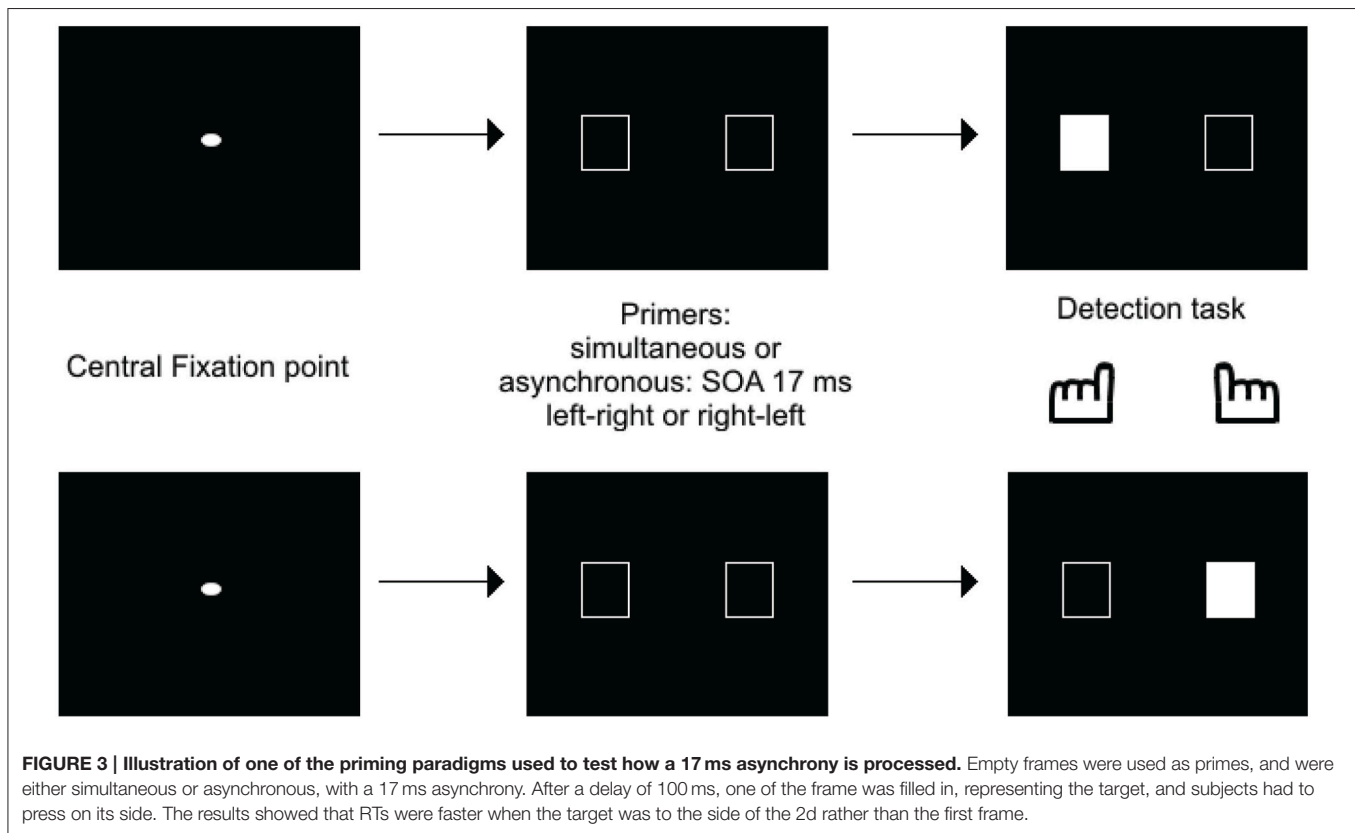


**FIGURE 2 | Illustration of the paradigm used to measure a Simon effect when stimuli are asynchronous.** The asynchrony is manipulated from 0 to 100 ms by 8 or 17 ms steps. The subjects decide whether the stimuli are simultaneous or asynchronous and press a response key accordingly. The Simon effect shows in a tendency to press on either the first or second stimulus side, whatever the side of this stimulus, as shown on the figure.

while inhibition would indicate that attention had shifted from the prime location. This procedure thus allowed examination of how attention shifts as a function of prime presentation. The time course of facilitation and/or inhibition was studied by examining how either evolved over a range of delays between primes and targets. Importantly, asynchronies below 20 ms were non-detectable, and these asynchronies should have led stimuli to be integrated within the same temporal window. Yet, the results confirmed that stimuli presented with asynchronies of 17 ms were processed as temporally separate events. The results also suggested that successive primes were processed serially, consistent with an attentional account. This was indicated by inhibition on the side of the first prime after a short delay (50 ms between primes and target) and by the facilitation on the side of the second prime (100 ms after the occurrence of primes). Importantly, we checked that these effects did not depend on the side of the hand response by changing the response mode (answering on the side of the first vs. second target). Effects are rather a consequence of a shift of attention (see Poncelet and Giersch, 2015, for a more detailed discussion on alternative explanations).

Several studies have now established that visual stimuli are distinguished in time at an implicit level even when belonging to the same perceptual moment. Moreover, it seems healthy

observers can unconsciously follow events of the same temporal moment over time, possibly by displacing their attention from one event to the other (Poncelet and Giersch, 2015). All in all these results suggest that information processing is temporally structured even within perceptual moments. It might seem surprising that these effects stayed unnoticed, but this might have been so because most paradigms involve integration processes. For example the flash-lag effect relies on the display of a moving object and a flashing light at some point of the trajectory. Typically, the moving object is perceived ahead of its real location at the time of the flashing light. The shift can correspond to a delay as long as a perceptual moment (25–45 ms; Whitney and Murakami, 1998; Kanai et al., 2004). The implicit processing of information in time should lead to higher precision, but seems not to prevent illusions to occur. This means it is the perceptual moment that shapes conscious experience, even in tasks that do not require an explicit temporal judgment. As a consequence the conscious experience may mask the influence of unconscious mechanisms operating over short delays, i.e., within the perceptual moment. It does not mean, however, that such unconscious mechanisms have no impact on our conscious experience. It only means that this influence is obscured by the operation of other processes, such as postdiction mechanisms (Shimojo, 2014). These help to interpret and render information



in the environment as meaningful. In the real world our sensory systems are continuously subjected to multiple, unrelated signals. Under these circumstances, automatically integrating successive events within temporal windows would not help to make sense of this information, while an additional processing step might prove helpful. The availability of individual events before their integration within a perceptual moment, together with a progressive displacement of attention following the events, may be used to apply filters and choose to which extent events will be included within the perceptual moment. This might explain the length of the integration process as described in Scharnowski et al. (2009) and Pilz et al. (2013). Such an hypothesis is speculative and requires confirmation. However, what is clear is that the processing of perceptual information is refreshed at high frequency and that the integration and fusion of information does not preclude access to individual and successive information within perceptual moments.

The high frequency of the information processing refreshment rate converges with the results of Elliott et al. (2007), who has shown that sub-threshold synchrony up to 14–21 ms primes the detection of simultaneous targets. This interval, as suggested above, may correspond to the time required to establish neuronal assemblies of synchronized neurons, and thus to integrate information. This kind of integration, however, would mainly correspond to the binding of single events in time: in this case the presentation of two synchronous or quasi-synchronous stimuli. With reference to the binding of discrete neural processes in

ever-changing event structures, we might expand definition of “event structure” to include all events, including the neural events to which the responses of any two functionally separable processes would have to respond. So for example, different neural assemblies are responsible for coding the color and the direction of motion of an object, and their binding ensures the moving item maintains object constancy (i.e., it is perceived as the same object and the same color) in spite of the movement of the object and its spatial displacement, as well as factors such as the observers eye movements (which might include micro-tremor and fast, stimulus-independent oscillations, e.g., Neuenschwander and Singer, 1996). On this basis, relatively fast binding, operating at elementary levels of perceptual processing might be necessary to ensure correct bindings are coded and maintained in spite of unpredictable changes in the event structure to which the synchronized assembly responds. It is only when events are more distant in time, i.e., above 14–21 ms, that assemblies for each event would be distinguished from one another. Such assemblies would be local, and would allow for successive processing of events. However, extracting information on asynchrony and order may require additional processing entailing a comparison of the two events. It is this comparison that would then be accessed consciously, possibly based on longer-range synchronization phenomena. This might be one explanation for the fact that events that are 17 ms apart can be automatically and unconsciously followed in time, based on successive local synchronization phenomena, but that conscious



separation of events in time occurs at larger asynchronies only. This possibility is also consistent with the observations that some time is needed to relate successive events with one another, and to integrate them into conscious forms (Scharnowski et al., 2009; Pilz et al., 2013) and across perceptual moments.

The fact that information is automatically distinguished in time within intervals as short as 17 ms is not necessarily in contradiction with the concepts of temporal windows, inasmuch it mainly adds an additional, implicit level of processing. What requires consideration is how evidence for implicit processing changes our understanding of the emergence of the sense of

time continuity. The fact that we are able to process and follow information over time with high temporal fidelity seems to contribute to our feeling of time continuity. Indeed, the fact that we can process information with a better time accuracy at an unconscious than at a conscious level means that any environmental change between successive 50 ms windows can be resolved by means of a smoothing over processes responsible for event coding. In addition, each time we consciously look, potentially we unconsciously check for new information several times. In this sense, it can be approximated that we have a continuous access to the outer world.

## REFERENCES

- Allport, D. A. (1964). Phenomenal simultaneity and the perceptual moment hypothesis. *Br. J. Psychol.* 59, 395–406. doi: 10.1111/j.2044-8295.1968.tb01154.x
- Anderson, H. K., and Grush, R. (2009). A brief history of time-consciousness: historical precursors to James and Husserl. *J. Hist. Philos.* 47, 277–307. doi: 10.1353/hph.0.0118
- Babkoff, H., and Fostick, L. (2013). The role of tone duration in dichotic temporal order judgment. *Atten. Percept. Psychophys.* 75, 654–660. doi: 10.3758/s13414-013-0449-6
- Brecher, G. A. (1932). Die Entstehung und biologische Bedeutung der subjectiven Zeiteinheit—des Momentes. *Z. Vgl. Physiol.* 18, 204–243.
- Dainton, B. (2010). “Temporal consciousness,” in *The Stanford Encyclopedia of Philosophy*, ed E. N. Zalta. Available online at: <http://plato.stanford.edu/archives/fall2010/entries/consciousness-temporal>
- Duncan, J., and Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychol. Rev.* 96, 433–458. doi: 10.1037/0033-295X.96.3.433
- Efron, R. (1970a). The relationship between the duration of a stimulus and the duration of a perception. *Neuropsychologia* 8, 37–55. doi: 10.1016/0028-3932(70)90024-2
- Efron, R. (1970b). The minimum duration of a perception. *Neuropsychologia* 8, 57–63. doi: 10.1016/0028-3932(70)90025-4
- Elliott, M. A., Shi, Z., and Surer, F. (2007). The effects of subthreshold synchrony on the perception of simultaneity. *Psychol. Res.* 71, 687–693. doi: 10.1007/s00426-006-0057-3
- Exner, S. (1875). Experimentelle Untersuchungen der einfachsten psychischen Prozesse. *Pflügers Arch.* 11, 403–432. doi: 10.1007/BF01659311
- Fink, M., Ulbrich, P., Churan, J., and Wittmann, M. (2006). Stimulus-dependent processing of temporal order. *Behav. Process.* 71, 344–352. doi: 10.1016/j.beproc.2005.12.007
- Foucher, J. R., Lacambre, M., Pham, B. T., Giersch, A., and Elliott, M. A. (2007). Low time resolution in schizophrenia lengthened windows of simultaneity for visual, auditory and bimodal stimuli. *Schizophr. Res.* 97, 118–127. doi: 10.1016/j.schres.2007.08.013
- Fuchs, T. (2007). The temporal structure of intentionality and its disturbance in schizophrenia. *Psychopathology* 40, 229–235. doi: 10.1159/000101365
- Gepshtein, S., and Kubovy, M. (2000). The emergence of visual objects in space-time. *Proc. Natl. Acad. Sci. U.S.A.* 97, 8186–8191. doi: 10.1073/pnas.97.14.8186
- Giersch, A., Lalanne, L., Corves, C., Seubert, J., Shi, Z., Foucher, J., et al. (2009). Extended visual simultaneity thresholds in patients with schizophrenia. *Schizophr. Bull.* 35, 816–825. doi: 10.1093/schbul/sbn016
- Giersch, A., Lalanne, L., van Assche, M., and Elliott, M. E. (2013). On disturbed time continuity in schizophrenia: an elementary impairment in visual perception? *Front. Psychol.* 4:281. doi: 10.3389/fpsyg.2013.00281
- Giersch, A., Poncelet, P., Capa, R. L., Martin, B., Duval, C. Z., Curziatti, M., et al. (2015). Disruption of information processing in schizophrenia: the time perspective. *Schizophr. Res. Cogn.* 2, 78–83. doi: 10.1016/j.scog.2015.04.002
- Gray, C. M., König, P., Engel, A. K., and Singer, W. (1989). Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature* 338, 334–337. doi: 10.1038/338334a0
- Hommel, B. (2011a). The Simon effect as a tool and heuristic. *Acta Psychol.* 136, 189–202. doi: 10.1016/j.actpsy.2010.04.011
- Hommel, B. (2011b). Attention and spatial stimulus coding in the Simon task: are joiner to van der Lubbe and Abrahamse (2010). *Acta Psychol.* 136, 265–268. doi: 10.1016/j.actpsy.2010.10.002
- Kanai, R., Sheth, B. R., and Shimojo, S. (2004). Stopping the motion and sleuthing the flash-lag effect: spatial uncertainty is the key to perceptual mislocalization. *Vis. Res.* 44, 2605–2619. doi: 10.1016/j.visres.2003.10.028
- Lalanne, L. (1876). Sur la duree de la sensation tactile. *C. R. Acad. Sci. Paris* 90, 1314–1316.
- Lalanne, L., Van Assche, M., and Giersch, A. (2012a). When predictive mechanisms go wrong: disordered visual synchrony thresholds in schizophrenia. *Schizophr. Bull.* 38, 506–513. doi: 10.1093/schbul/sbq107
- Lalanne, L., van Assche, M., Wang, W., and Giersch, A. (2012b). Looking forward: an impaired ability in patients with schizophrenia? *Neuropsychologia* 50, 2736–2744. doi: 10.1016/j.neuropsychologia.2012.07.023
- Martin, M., Giersch, A., Huron, C., and van Wassenhove, V. (2013). Temporal event structure and timing in schizophrenia: preserved binding in a longer “now” *Neuropsychologia* 51, 358–371. doi: 10.1016/j.neuropsychologia.2012.07.002
- Martin, B., Wittmann, M., Franck, N., Cermolacce, M., Berna, F., and Giersch, A. (2014). Temporal structure of consciousness and minimal self in schizophrenia. *Front. Psychol.* 5:1175. doi: 10.3389/fpsyg.2014.01175
- Neisser, U. (1967). *Cognitive Psychology*. New York: Appleton Century Crofts.
- Neuenschwander, S., and Singer, W. (1996). Long range synchronisation of oscillatory light responses in the cat retina and lateral geniculate nucleus. *Nature* 379, 728–733. doi: 10.1038/379728a0
- Phillips, I. (2014). “The temporal structure of experience,” in *Subjective Time: The Philosophy, Psychology and Neuroscience of Temporality*, eds V. Arstila and D. Lloyd (Cambridge MA: MIT Press), 139–158.
- Pilz, K. S., Zimmermann, C., Scholz, J., and Herzog, M. H. (2013). Long-lasting visual integration of form, motion, and color as revealed by visual masking. *J. Vis.* 13, 1–11. doi: 10.1167/13.10.12
- Poncelet, P. E., and Giersch, A. (2015). Tracking visual events in time in the absence of time perception: implicit processing at the ms level. *PLoS ONE* 10:e0127106. doi: 10.1371/journal.pone.0127106
- Pöppel, E. (1997). A hierarchical model of temporal perception. *Trends Cogn. Sci.* 1, 56–61. doi: 10.1016/S1364-6613(97)01008-5
- Scharnowski, F., Rüter, J., Jolij, J., Hermens, F., Kammer, T., and Herzog, M. H. (2009). Long-lasting modulation of feature integration by transcranial magnetic stimulation. *J. Vis.* 9, 1–10. doi: 10.1167/9.6.1
- Schmidt, H., McFarland, J., Ahmed, M., McDonald, C., and Elliott, M. A. (2011). Low-level temporal coding impairments in psychosis: preliminary findings and recommendations for further studies. *J. Abnorm. Psychol.* 120, 476–482. doi: 10.1037/a0023387
- Shimojo, S. (2014). Postdiction: its implications on visual awareness, hindsight, and sense of agency. *Front. Psychol.* 5:196. doi: 10.3389/fpsyg.2014.00196
- Simon, J. R. (1969). Reactions towards the source of stimulation. *J. Exp. Psychol.* 81, 174–176. doi: 10.1037/h0027448

- Singer, W. (1993). Synchronization of cortical activity and its putative role in information processing and learning. *Annu. Rev. Physiol.* 55, 349–374. doi: 10.1146/annurev.ph.55.030193.002025
- Singer, W. (1999). Neuronal synchrony: a versatile code for the definition of relations? *Neuron* 24, 49–65. doi: 10.1016/S0896-6273(00)80821-1
- Stroud, J. M. (1955). “The fine structure of psychological time,” in *Information Theory in Psychology: Problems and Methods* ed. H. Quastler (Glencoe, IL: Free Press), 174–207.
- Sweet, A. L. (1953). Temporal discrimination by the human eye. *Am. J. Psychol.* 66, 185–198. doi: 10.2307/1418725
- Ullman, S. (1979). *The Interpretation of Visual Motion*. Cambridge, MA: MIT Press.
- van der Lubbe, R. H. J., and Abrahamse, E. L. (2011). The premotor theory of attention and the Simon effect. *Acta Psychol.* 136, 259–264. doi: 10.1016/j.actpsy.2010.09.007
- van Rullen, E., Zoefel, B., and Ilhan, B. (2014). On the cyclic nature of perception in vision versus audition. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 369, 20130214. doi: 10.1098/rstb.2013.0214
- van Wassenhove, V. (2009). Minding time in an amodal representational space. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 1815–1830. doi: 10.1098/rstb.2009.0023
- Vatakis, A., and Spence, C. (2007). Crossmodal binding: evaluating the “unity assumption” using audiovisual speech stimuli. *Percept. Psychophys.* 69, 744–756. doi: 10.3758/BF03193776
- Von Baer, K. E. (ed.). (1864). “Welche Auffassung der lebenden Natur ist die richtige? Und wie ist diese Auffassung auf die Entomologie anzuwenden?” in *Reden, gehalten in wissenschaftlichen Versammlungen und kleinere Aufsätze vermischten Inhalt, (the first part given as an oral lecture in 1860)* (St. Petersburg: H. Schmitzdorf), 237–284.
- Von Baer, K. E. (1909). Welche Auffassung der lebenden Natur ist die richtige? Und wie ist diese Auffassung auf die Entomologie anzuwenden? *Aus baltischer Geistesarbeit: Reden und Aufsätze* 1, 1–47.
- von Uexküll, J. (1957). “A stroll through the worlds of animals and men: a picture book of invisible worlds,” *Instinctive Behavior: The Development of a Modern Concept*, ed. and trans. C. H. Schiller (New York: International Universities Press, Inc.), 5–80.
- Wehrhahn, C., and Rapf, D. (1992). ON- and OFF-pathways form separate neural substrates for motion perception: psychophysical evidence. *J. Neurosci.* 12, 2247–2250.
- Westheimer, G., and McKee, S. P. (1977). Perception of temporal order in adjacent visual stimuli. *Vis. Res.* 17, 887–892. doi: 10.1016/0042-6989(77)90062-1
- Whitney, D., and Murakami, I. (1998). Latency difference, not spatial extrapolation. *Nat. Neurosci.* 1, 656–657. doi: 10.1038/3659
- Wittmann, M. (2011). Moments in time. *Front. Integr. Neurosci.* 5:66. doi: 10.3389/fnint.2011.00066

**Conflict of Interest Statement:** 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.

Copyright © 2016 Elliott and Giersch. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Defense of the brain time view

Valtteri Arstila<sup>1,2\*</sup>

<sup>1</sup> Department of Behavioral Sciences and Philosophy, University of Turku, Turku, Finland, <sup>2</sup> Turku Brain and Mind Center, University of Turku, Turku, Finland

**Keywords:** brain time view, time-marker view, perceptual asynchrony, temporal judgment, reaction time

## Introduction

Making sense of our ever-changing surroundings requires the extraction of temporal information from a continuous stream of stimulation. How do we achieve this when the relevant information can be separated by only tens of milliseconds? I address this question by focusing on the debate concerning the perceptual asynchrony of changes of visual features. This refers to the finding that when a moving stimulus simultaneously changes its color and direction of movement, the subjects report that the change in color occurs 60–100 ms before the change in direction.

One explanation for this finding is that colors are processed faster than motion, which in turn means that changes in color are processed faster than changes in motion, and this difference in perceptual latency is reflected in our judgments of the temporal features of the stimuli (Moutoussis and Zeki, 1997, 2002). Crucially, this explanation assumes that the judged order of events mirrors the time at which the brain generates the representation of the events or their features. Because the temporal properties of the representations generated by the brain serve as time-markers<sup>1</sup>, this view has been called *the brain time view*.

Nishida and Johnston (2002, 2010) object to this explanation and propose *the time-marker view* as an alternative. This view differs from the brain time view in two respects. First, it holds that representations of color and motion are generated at the same time, and that the reported asynchrony between them results from an error in a specialized mechanism responsible for temporal judgments. Second, temporal judgments mirror the timing of external events as closely as possible (rather than the time when the neural processing of the events is completed). For this reason, the mechanism is thought to be a mid-level perceptual process.

In what follows, I will defend the brain time view from the objections raised against it by Nishida and Johnston. This has been already done on the empirical grounds (e.g., Arnold, 2010; Moutoussis, 2012, 2014). My argumentation complements this debate by focusing on the more theoretical issues and highlighting implicit assumptions in Nishida and Johnston's argumentation.

## The Inherent Problems of the Brain Time View

The first set of objections concerns a number of inherent problems that the brain time view allegedly faces. To begin with, referring to Dennett and Kinsbourne (1992), Johnston and Nishida (2001, R428) argue that the brain time view faces "some thorny philosophical problems." Yet, not all of them are particularly pressing. For example, Nishida and Johnston (2002) argue that the brain time view comprises "a logical pitfall" because it equates the time when the event appears to occur with the time when the brain generates the representation. Even though it is theoretically possible that these two can be separated, it does not follow that they actually are. Thus, pointing out the possibility is not a particularly effective objection.

Two inherent problems need to be addressed in more detail, however. First, Nishida and Johnston (2010, 286) claim that the brain time view suffers the "logical shortcoming of identifying

<sup>1</sup>In its most general sense, a time-marker is something that a process can utilize during a task in which the temporal features of stimuli or experiences of stimuli are determined.

## OPEN ACCESS

### Edited by:

Marc Wittmann,  
Institute for Frontier Areas of  
Psychology and Mental Health,  
Germany

### Reviewed by:

Kielan Yarrow,  
City University London, UK

### \*Correspondence:

Valtteri Arstila,  
valtteri.arstila@utu.fi

### Specialty section:

This article was submitted to  
Perception Science,  
a section of the journal  
Frontiers in Psychology

**Received:** 05 July 2015

**Accepted:** 24 August 2015

**Published:** 04 September 2015

### Citation:

Arstila V (2015) Defense of the brain  
time view. *Front. Psychol.* 6:1350.  
doi: 10.3389/fpsyg.2015.01350

physical co-occurrence of the cortical representation of event A and that of event B with the representation of co-occurrence of events A and B.” Second, in relation to the brain time view, they (Nishida and Johnston, 2002, 359) claim that “[i]n order to judge the temporal order of two arbitrary neural events, the brain must have a mechanism to compare them and anatomical connections of high temporal fidelity between the neurons to be compared. This meta-analysis of neural processing places a high combinatorial burden on the brain.” These two claims are inconsistent though. While the first holds that the co-occurrence of two neural events is not separately represented, the second argues that the temporal judgments are due to some sort of comparison mechanism. The second claim is a more plausible option and concurs with, say, Efron’s (1963) idea of simultaneity center according to which the temporal order of stimuli is determined on the basis of the relative arrival times of sensory signals to this hypothetical simultaneity center. Thus, the objection based on a logical shortcoming is misdirected.

As for the claim concerning a burden on the brain, it too is misdirected as it assumes that the purpose of the brain time view is to deduce the actual time of external events from the time of neural activity—this is the reason why the temporal comparator supposedly needs to account for the temporal features of different anatomical connections. However, such an assumption is not subscribed to by proponents of the brain time view, nor is it part of the brain time view as Nishida and Johnston themselves describe it! On the contrary, the brain time view holds that timing mirrors when the representations of events are generated. (Mirroring not mean that timing needs to match exactly with the time when the representations of events are generated.) Thus, the mechanism needs to compare only the temporal properties of neural signals—just as Efron’s simultaneity center does—and, since this task is necessitated by the time-marker view as well, the burden on the brain is the same in this respect.

It is worth highlighting another commonality between these two views: in both theories, time-marker is a *temporal property of neural activity brought about by an external event that some timing mechanism makes use of*<sup>2</sup>. For example, in the time-marker view, such a mechanism makes use of “the temporal pattern of the neural activity elicited by [external] events.” (Nishida and Johnston, 2002, 366) Thus, the difference here is that in the time-marker view the time-marker is based on an unconscious, mid-level perceptual neural activity, whereas in the

brain time view the neural activity relates to the representation of the event, and the temporal mechanism comes into play later in the processing hierarchy.

## The Brain Time View and Inconsistent Latency Estimations

The second objection against the brain time view is based on inconsistencies in perceptual latency estimations obtained using the reaction time method and temporal judgment tasks<sup>3</sup>. Nishida and Johnston (2002, 362) claim that the inconsistent results are problematic for the brain time view because, in the context of the perceptual asynchrony debate, “it is difficult to understand why [the asynchrony measured with temporal judgments] is not reflected in reaction time.”

The described inconsistency assumes, however, that an external event produces only one time-marker and that assumption has been rejected in two equally reasonable ways. Sternberg and Knoll’s *different time-marker hypothesis* (1973) rejects the assumption by maintaining that the two tasks have different task demands: temporal order judgments maximize correct judgments, whereas reaction time tasks emphasize speed (Sternberg and Knoll, 1973). Miller and Schwarz (2006, 394) likewise maintain that these two tasks have “fundamentally different task demands.” Thus, the tasks use different features of a single internal response as time-markers<sup>4</sup> and subsequently produce different results based on the same response. This concurs with the previous notion of time-markers because the constitution of a time-marker depends on the timing mechanisms and, hence, one internal response can manifest as different time-markers if different timing mechanisms make use of different temporal properties of the response.

Another way to reject the assumption is to argue that an external event causes two different internal responses, both of which serve as time-markers (Tappe et al., 1994; Aschersleben and Müsseler, 1999). One response is utilized by temporal order judgments and occurs in the later stage of processing. Early on, the processing leading to this response is separated from the processing which feeds into the motor system and is used in the reaction time tasks. Because, there are two timing mechanisms that use different internal responses as the basis for time-markers, the two mechanisms can provide inconsistent results. In both alternatives, temporal judgments make use of temporal properties of neural states, which can be representations of events, and thus the brain time view can account for the inconsistency.

<sup>2</sup>This assumes that the outcome of the mid-level mechanism responsible for temporal judgments in the time-marker view is something that we become conscious of, rather than something that is utilized by some subsequent timing mechanism. Theories that make use of this kind notion of time-markers often concern simultaneity perception and reaction time studies (e.g., Efron, 1963; Jaśkowski, 1996; Jaśkowski et al., 2014; Yarrow and Arnold, 2015). This notion can be contrasted with a symbolic notion that is often attributed to Dennett and Kinsbourne (1992). They illustrated it using date stamps on letters—a stamp (time-marker) represents the date when a letter is sent regardless of when the stamp is interpreted (when the letter arrives). This notion is rarely explicitly endorsed—something along these lines has been presented in relation to the postdiction effects (e.g., Eagleman and Sejnowski, 2000, 2007; Grush, 2005, 2006)—and even more rarely explicated. Thus, the notion remains under-described, both theoretically and in neural terms (e.g., Arnold, 2010; Arstila, 2015).

<sup>3</sup>For example, reaction time measurements and temporal order judgments are affected differently by changes in stimulus intensity and its luminance profile (Roufs, 1974; Jaśkowski, 1996), stimulus modality (Rutschmann and Link, 1964; Jaśkowski et al., 1990), spatial frequency of visual gratings (Tappe et al., 1994), and stimulus motion (Aschersleben and Müsseler, 1999).

<sup>4</sup>According to Sternberg and Knoll, TOJ tasks use the time of the activation peak of the response and RT tasks use the time when the activation crosses some earlier threshold. Miller and Schwarz, in turn, argue that the criterion in RT tasks is higher than in TOJ tasks, and thus that the RT tasks employ the later part of the internal response than TOJ tasks.



## Latencies of Different Types of Changes

Finally, Nishida and Johnston (2002, 2010) make a distinction between first-order and second-order changes. The first-order changes, called *transitions*, require that two points in time be compared (e.g., color at  $t_1$  and  $t_2$ ). Second-order changes, called *turning points*, require the comparing of three points in time (e.g., spatial position at  $t_1$ ,  $t_2$ , and  $t_3$ ). A special mechanism is thought to exist only for temporal judgments related to transitions and thus determining the time of turning points takes longer. Consequently, even if a transition and a turning point were to occur simultaneously, the latter would be judged to occur later than the former. Nishida and Johnston's results as regards synchrony between different transitions and turning points support this claim.

Assuming that these experiments are comparable to those concerning the original finding, the obtained results conflict with the claim that colors are processed faster than motion. However, they are not in conflict with the brain time view in general. This is because the results do not specify the processing stage at which the time-markers for the turning points are established. Hence, they are also compatible with the claim that the mechanism responsible for temporal judgments in the brain time view requires more time to determine turning points

than to determine transitions. In this way, the brain time view can explain Nishida and Johnston's finding in largely the same fashion as the time-marker view.

## Summary

The brain time view and the time-marker view can be understood to rely upon the existence of a temporal judgment comparator that makes use of the temporal properties of neural activity caused by an external event. The main difference in these two views concerns the processing stage in which such comparison takes place and what the timing concerns about. The objections raised by Nishida and Johnston against the brain time view cannot settle the question of which theory is closer to the truth. However, Nishida and Johnston (2010, 286) are correct in their claim that the brain time view "assumes that a brain time mechanism is poorly designed in the sense that processing delay is added to event time estimation." Then again, given the evidence that cortical processing influences temporal judgments (Arnold and Wilcock, 2007), and that Efron (1963) postulated his simultaneity center exactly because it could account for the processing delays between cortical hemispheres, the existence of such a poor mechanism could be closer to the truth than the mechanism postulated by the time-marker view.

## References

- Arnold, D. H., and Wilcock, P. (2007). Cortical processing and perceived timing. *Proc. R. Soc. B Biol. Sci.* 274, 2331–2336. doi: 10.1098/rspb.2007.0576
- Arnold, D. H. (2010). "Relative timing and perceptual asynchrony," in *Space and Time in Perception and Action*, eds R. Nijhawan and B. Khurana (Cambridge, UK: Cambridge University Press), 254–277.
- Arstila, V. (2015). "The time of experience and the experience of time," in *Philosophy and Psychology of Time*, eds B. Mölder, V. Arstila, and P. Øhrstrøm (Dordrecht: Springer).
- Aschersleben, G., and Müsseler, J. (1999). Dissociations in the timing of stationary and moving stimuli. *J. Exp. Psychol. Hum. Percept. Perform.* 25, 1709–1720. doi: 10.1037/0096-1523.25.6.1709
- Dennett, D. C., and Kinsbourne, M. (1992). Time and the observer. *Behav. Brain Sci.* 15, 183–247. doi: 10.1017/S0140525X00068229
- Eagleman, D. M., and Sejnowski, T. J. (2000). Motion integration and postdiction in visual awareness. *Science* 287, 2036–2038.
- Eagleman, D. M., and Sejnowski, T. J. (2007). Motion signals bias localization judgments: a unified explanation for the flash-lag, flash-drag, flash-jump, and Frohlich illusions. *J. Vis.* 7:3. doi: 10.1167/7.4.3
- Efron, R. (1963). The effect of handedness on the perception of simultaneity and temporal order. *Brain* 86, 261–284. doi: 10.1093/brain/86.2.261
- Grush, R. (2005). "Brain time and phenomenological time," in *Cognition and the Brain: The Philosophy and Neuroscience Movement*, eds A. Brook and K. Akins (Cambridge, MA: Cambridge University Press), 160–207.
- Grush, R. (2006). How to, and how not to, bridge computational cognitive neuroscience and Husserlian phenomenology of time consciousness. *Synthese* 153, 417–450. doi: 10.1007/s11229-006-9100-6
- Jaśkowski, P. (1996). Simple reaction time and perception of temporal order: dissociations and hypotheses. *Percept. Mot. Skills* 82, 707–730. doi: 10.2466/pms.1996.82.3.707
- Jaśkowski, P. (2014). "What determines simultaneity and order perception?," in *Subjective Time: The Philosophy, Psychology, and Neuroscience of Temporality*, eds V. Arstila and D. Lloyd (Cambridge, MA: MIT Press), 379–408.
- Jaśkowski, P., Rybarczyk, K., Jaroszyk, F., and Hojan-Jezińska, D. (1990). Temporal-order judgments and reaction time for stimuli of different modalities. *Psychol. Res.* 52, 35–38. doi: 10.1007/BF00867209
- Johnston, A., and Nishida, S. (2001). Time perception: brain time or event time? *Curr. Biol.* 11, R427–R430. doi: 10.1016/S0960-9822(01)00252-4
- Miller, J., and Schwarz, W. (2006). Dissociations between reaction times and temporal order judgments: a diffusion model approach. *J. Exp. Psychol. Hum. Percept. Perform.* 32, 394–412. doi: 10.1037/0096-1523.32.2.394
- Moutoussis, K. (2012). Asynchrony in visual consciousness and the possible involvement of attention. *Front. Psychol.* 3:314. doi: 10.3389/fpsyg.2012.00314
- Moutoussis, K. (2014). "Perceptual asynchrony in vision," in *Subjective Time: The Philosophy, Psychology, and Neuroscience of Temporality*, eds V. Arstila and D. Lloyd (Cambridge, MA: The MIT Press), 201–216.
- Moutoussis, K., and Zeki, S. (1997). A direct demonstration of perceptual asynchrony in vision. *Proc. R. Soc. B Biol. Sci.* 264, 393–399. doi: 10.1098/rspb.1997.0056
- Moutoussis, K., and Zeki, S. (2002). The relationship between cortical activation and perception investigated with invisible stimuli. *Proc. Natl. Acad. Sci. U.S.A.* 99, 9527–9532. doi: 10.1073/pnas.142305699
- Nishida, S., and Johnston, A. (2002). Marker correspondence, not processing latency, determines temporal binding of visual attributes. *Curr. Biol.* 12, 359–368. doi: 10.1016/S0960-9822(02)00698-X
- Nishida, S., and Johnston, A. (2010). "The time marker account of cross-channel temporal judgments," in *Space and Time in Perception and Action*, eds R. Nijhawan and B. Khurana (Cambridge, UK: Cambridge University Press), 278–300.
- Roufs, J. (1974). Dynamic properties of vision—V: perception lag and reaction time in relation to flicker and flash thresholds. *Vision Res.* 14, 853–869. doi: 10.1016/0042-6989(74)90149-7
- Rutschmann, J., and Link, R. (1964). Perception of temporal order of stimuli differing in sense mode and simple reaction time. *Percept. Mot. Skills* 18, 345–352. doi: 10.2466/pms.1964.18.2.345

- Sternberg, S., and Knoll, R. L. (1973). "The Perception of temporal order: fundamental issues and a general model," in *Attention and Performance IV*, ed S. Kornblum (New York, NY: Academic Press), 629–685.
- Tappe, T., Niepel, M., and Neumann, O. (1994). A dissociation between reaction time to sinusoidal gratings and temporal-order judgment. *Perception* 23, 335–347. doi: 10.1068/p230335
- Yarrow, K., and Arnold, D. H. (2015). "The timing of experiences: how far can we get with simple brain-time models?," in *Philosophy and Psychology of Time*, eds B. Mölder, V. Arstila, and P. Øhrstrøm (Dordrecht: Springer).

**Conflict of Interest Statement:** 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.

Copyright © 2015 Arstila. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Temporal perception in visual processing as a research tool

Bin Zhou<sup>1\*</sup>, Ting Zhang<sup>2</sup> and Lihua Mao<sup>2\*</sup>

<sup>1</sup> Key Laboratory of Mental Health, Institute of Psychology, Chinese Academy of Sciences, Beijing, China, <sup>2</sup> Department of Psychology, Peking University, Beijing, China

## OPEN ACCESS

### Edited by:

Marc Wittmann,  
Institute for Frontier Areas  
of Psychology and Mental Health,  
Germany

### Reviewed by:

Vani Pariyadath,  
National Institutes of Health/National  
Institute on Drug Abuse, USA  
Kentaro Yamamoto,  
University of Tokyo, Japan

### \*Correspondence:

Bin Zhou,  
Key Laboratory of Mental Health,  
Institute of Psychology, Chinese  
Academy of Sciences,  
Lincui Road 16,  
Beijing 100101, China  
zhoub@psych.ac.cn;  
Lihua Mao,  
Department of Psychology, Peking  
University, Yiheyuan Road 5,  
Beijing 100871, China  
maolihua@pku.edu.cn

### Specialty section:

This article was submitted to  
Perception Science,  
a section of the journal  
Frontiers in Psychology

**Received:** 22 February 2015

**Accepted:** 13 April 2015

**Published:** 24 April 2015

### Citation:

Zhou B, Zhang T and Mao L (2015)  
Temporal perception in visual  
processing as a research tool.  
Front. Psychol. 6:521.  
doi: 10.3389/fpsyg.2015.00521

Accumulated evidence has shown that the subjective time in the sub-second range can be altered by different factors; some are related to stimulus features such as luminance contrast and spatial frequency, others are processes like perceptual grouping and contextual modulation. These findings indicate that temporal perception uses neural signals involved in non-temporal feature processes and that perceptual organization plays an important role in shaping the experience of elapsed time. We suggest that the temporal representation of objects can be treated as a feature of objects. This new concept implies that psychological time can serve as a tool to study the principles of neural codes in the perception of objects like “reaction time (RT).” Whereas “RT” usually reflects the state of transient signals crossing decision thresholds, “apparent time” in addition reveals the dynamics of sustained signals, thus providing complementary information of what has been obtained from “RT” studies.

**Keywords:** subjective time, perceptual organization, research tool, neural response, object perception

## Introduction

One of the extensively studied temporal topics is duration or interval timing in the sub-second range, with which a basic experience beyond instantaneity can be obtained (Fraisse, 1984). Such temporal processing is distinguished from time estimation of supra-second intervals that involves higher-order cognitive strategies (Fraisse, 1984). Whereas perceptual signal integration across tens to hundreds milliseconds is assumed to be automatic and pre-semantic (Pöppel, 1997, 2009), the feeling of how long an event lasts on this timescale appears to engage some active cognitive processes such as memory operations. Apparent time and related phenomena are often retrospectively constructed, and the context in which the temporal pattern is embedded usually exerts a modulatory effect on the time that one experiences (Eagleman, 2008; Bao et al., 2013). This retrospective and context-sensitive characteristic has also been discussed by Zhou et al. (2014a) and may have some connections with previous opinions on the phenomenological feature of subjective time (Woodrow, 1951; Gibson, 1975). Given that the spatiotemporal context often determines how stimuli are perceptually organized and perceived (Kramer and Yantis, 1997; Wagemans et al., 2012), it is important to consider the perceptual organization and its relationship to subjective time when one tries to understand the mechanisms underpinning various temporal phenomena.

## Sub-Second Time is Modulated by Perceptual Organization

Traditionally, the temporal signature of an event is examined separately from the content of that event, with the assumption that time and object features are represented by independent neural substrates (e.g., Pöppel, 1997, 2009). This independence, however, may not be true, considering that

temporal information is most likely decoded from neural signals which process object features. The spiking patterns and the signal-to-noise ratio of neural activities involved in object perception thus may have an influence on the temporal decoding processes. This argument is consistent with the view that the encoding of object-identity is one of the fundamental requirements to unify different temporal phenomena (Zhou et al., 2014a). Relevant opinions are also expressed as the hypothesis that neural coding efficiency offers a basis of subjective time (Eagleman and Pariyadath, 2009). More detailed models suggest that subjective time is the output of multi-stage processes with context dependent circuits interacting with the core temporal machinery (Merchant et al., 2013). In our view, the context dependent stages include object perception processes and provide feeding signals for the computation that may function at the central stage. What is perceptually represented (i.e., the content) may alter the interaction and thus modulate how the signals are decoded.

Previous studies have accumulated evidence supporting the association of object processes and perceived time; for example, apparent duration of a visual stimulus is modulated by a preceding prime stimulus (Zhou et al., 2010). When the stimulus-onset-asynchrony (SOA) is less than 40 ms, the prime (like a disk) is likely masked by and integrated with the target (a ring with inner diameter matching the diameter of the disk), leading to expanded duration of the target. However, when the SOA increases, the prime is consciously perceived and segmented from the target, resulting in duration compression. It appears that the perceptual organization between successive stimuli distorts the apparent duration of the stimulus. A similar illusion, in a form of chronostasis, can be demonstrated in situations where a voluntary action, such as a saccadic (Yarrow et al., 2001, 2004) or a manual (Park et al., 2003; Yarrow and Rothwell, 2003) movement, triggers the onset of a stimulus whose subjective duration is ultimately expanded relative to succeeding stimuli with the same physical durations. It is assumed that the illusory duration is substantiated by a mechanism to solve the uncertainty about the sensory onset caused by movement. However, other observations show the generalization of a chronostasis-like effect in audition and vision without voluntary action (Hodinott-Hill et al., 2002; Alexander et al., 2005; Hunt et al., 2008), indicating that motor action is neither necessary nor sufficient to produce chronostasis. Instead, more general functions such as arousal level, attention, and memory are proposed to account for the induced temporal dilation. Although these notions are intriguing, they may not cover the entire picture. A likely alternative is that both action-triggered and featural changes modulate the perceptual organization of the stimulus sequence that ultimately shape the object perception and the perceived duration of the target. Observers usually compare the relative duration of the target with that of other stimuli in the train, and it is arguable that the neural system segments the stimulus train and then estimates the target duration based on the average of comparison stimuli. Indeed, some studies have shown that visual temporal averaging (Ohya and Watanabe, 2007) and auditory rhythmic grouping (Thorpe and Trehub, 1989; Geiser and Gabrieli, 2013) alter the temporal structure of stimulus patterns and lead to relatively longer subjective intervals between compared to within perceptual groups. Our own studies with

double stimuli also reveal the role of segmentation and grouping of two visual events on the apparent duration of the second event (Zhou et al., 2010, 2014b). Interestingly, the subjective duration is modulated by the dominant grouping principle (Zhou et al., 2014b), thus manifests different patterns under various perceptual contexts.

It should be noted that the perceptual organization account does not conflict with other explanations of time distortion for brief intervals and durations. As a fundamental component of neural functions, perceptual organization profoundly interacts with other cognitive operations such as attention (Han et al., 2005; de Haan and Rorden, 2010), memory, and prior experience (Kimchi and Hadad, 2002; Zemel et al., 2002; Peterson and Berryhill, 2013). In a sense, perceptual organization under various influences sets the quality and content of our perceptual experiences (Herzog and Fahle, 2002; Wagemans et al., 2012), which may further serve as signals for the appreciation of elapsed time. On the basis of functional taxonomy (Pöppel, 1989), perceptual organization and temporal processing are logistical functions, whereas percepts represent content functions. Logistical functions usually modulate the way of contents being organized in content functions. There seems a loop that logistical functions influence content functions, which in turn carve the phenomenal representation of the former (Bao et al., 2013). Similar loop effects can be applied between logistical functions and other content functions, e.g., between the temporal organization and the content of memories. How does such loop implement itself using the “neural language”? A hypothesis advanced by Eagleman and Pariyadath (2009) may provide some clues; they argue that the consumed neural energy signals the length of the subjective duration, and neural repetition suppression is most probably underlying the time compression of repeated events. An oddball (Tse et al., 2004; Pariyadath and Eagleman, 2007, 2012; New and Scholl, 2009) or behaviorally salient stimulus (van Wassenhove et al., 2008; New and Scholl, 2009; Wittmann et al., 2010) can somehow escape the repetition suppression, thus leading to dilated apparent duration (relative to repeated stimuli). On the functional level, “oddball effects” may result from stimulus regularity and predictability (Pariyadath and Eagleman, 2007; Schindel et al., 2011) which nevertheless control the way of stimuli being perceptually organized. At what extent a stimulus can be segmented from other stimuli depends on its perceptual relation with others, on its ecological significance, or on its position in the stimulus stream. The finding that the duration of oddballs (Tse et al., 2004; Pariyadath and Eagleman, 2007, 2012), looming events (van Wassenhove et al., 2008; New and Scholl, 2009; Wittmann et al., 2010), and the first or last stimulus of a stream (Rose and Summers, 1995) is often overestimated may depend on their efficiency being separated from background events. The consequence of such process is the altered neural representation of these objects, and on the basis of the neural energy hypothesis (Eagleman and Pariyadath, 2009) their apparent durations. Similar within-sequence perceptual context and organization may also mediate the duration effect caused by speed changes even when the average speed is stable (Matthews, 2011a; Sasaki et al., 2013). Focusing on the relationship between subjective duration and perceptual organization highlights the importance of heuristic contexts and retrospective analyses on



tracking events' time. Thus, not only immediate contexts but also task environments can modulate the apparent duration of a stimulus (Pariyadath and Eagleman, 2007; Jazayeri and Shadlen, 2010; Matthews, 2011b; Zhou et al., 2014b).

As discussed above, perceptual organization modulates subjective duration via its role on shaping the representation of objects; it is natural to ask whether object features are also able to modify the perceived lapse of time. Answering this question, the apparent duration of a stimulus has been reported to link with its size (Xuan et al., 2007; Ono and Kitazawa, 2009; Alards-Tomalín et al., 2014), intensity (Nisly and Wasserman, 1989; Xuan et al., 2007), number of elements (Xuan et al., 2007), and spatial frequency (Aaen-Stockdale et al., 2011). Interestingly, psychological time is associated with perceived qualities and quantities of objects (Ono and Kawahara, 2007; Matthews et al., 2011; Yamamoto and Miura, 2012), rather than their physical properties. For example, a central circle appears to last longer when it is surrounded by smaller than by larger circles, although it has physically identical size and duration in both conditions (Ono and Kawahara, 2007). This effect is attributed to the different apparent sizes of the central circle induced by surrounding circles, a phenomenon termed Ebbinghaus illusion. In another study (Orgs et al., 2011), viewing an implied body, but not non-body or inverted-body, motion compresses subjective time. Interestingly, the compression effect depends on the apparent length of movement paths. Taken together with the observations in oddball and repetition paradigms, these findings strongly suggest that perceptual organization operating on both spatial and temporal domains influences how the target is temporally represented relative to other objects. This view exploits factors that shape the mental time from the point of perceptual functions and provides a useful supplement to other more biological concepts and hypotheses advocated previously (Eagleman and Pariyadath, 2009; Zhou et al., 2014a). Considering its dependence on neural signals encoding objects and their identities as well as its behavioral correlation with perceived object properties, subjective time can be treated as an object feature that is indirectly computed but nevertheless can be used to tag the object identity.

## Subjective Time Can Serve as a Research Tool to Examine Neural Representations of Objects

The association between the perceptual representation of objects and perceived time implies an important application of subjective time in the assessment of various mental processes. One can refer to the broad use of reaction time (RT) and its success in revealing characteristics of a variety of perceptual and cognitive operations since the beginning of experimental psychology (Pöppel, 1997; MacDonald and Meck, 2004; Posner, 2005). For a response to occur, accumulated neural signals, either detection or discrimination information, should reach a decision criterion which varies across trials but on average keeps stable for certain task sessions (Grice et al., 1982; Usher and McClelland, 2001). Usually, more strongly perceived objects lead to shorter RTs (Donner and Fagerholm, 2003; Palmer et al., 2005). Therefore, by inspecting RTs, one

may obtain rich information about the neural codes of external inputs or the cognitive structures processing these inputs. In a similar vein, one can compare apparent durations of different objects under the same condition or of the same object under different conditions to study object representation and related neural processes. For example, Zhou et al. (2014b) presented observers two successive Gabor patches with varying orientation differences or spatial distances across trials. The apparent duration of the second Gabor patch is underestimated relative to its physical duration. Interestingly, the amount of duration compression is positively related to the size of orientation difference and spatial distance, in a way consistent with the orientation tuning and retinotopic mapping of early visual neurons. Similar results are also reported by Pariyadath and Eagleman (2012) when using an oddball paradigm. Thus, properties of the visual cortex can be inferred from the perceived duration of a visual stimulus without adaptation manipulation or neurophysiological recordings. The application of this concept can be found in another study (Aaen-Stockdale et al., 2011). In a typical oddball paradigm, observers compared the relative duration of an infrequent oddball (a Gabor patch) and that of a frequent standard which could be either another Gabor patch with different spatial frequency or an auditory stimulus. The authors found an interesting result that the duration expansion or compression of the oddball depended on its spatial frequency. Relative to lower and higher spatial frequencies, a mid-range spatial frequency (ca. 2 c/deg) consistently led to longer apparent duration that was invariant across different baseline durations. The perceived time in this example thus indicates that longer visual persistence is associated with neural responses selective to mid-range spatial frequencies in the early visual cortices. Other reports demonstrate spatially localized duration effects for drifting gratings and suggest neural sources in visual areas of striate and dorsal extrastriate cortices where neurons have specific receptive fields for moving objects (Johnston et al., 2006; Burr et al., 2007; Bruno et al., 2010). For more complex events, such as biologically relevant stimuli, studying apparent time also helps to uncover the neural processes involved in encoding these events (Wang and Jiang, 2012; Yamamoto and Miura, 2012). It has been found that merely observing the movement of another organism activates the mirror neuron system of the primate brain (Rizzolatti and Craighero, 2004). The question is whether a static image with implied motion also activates mirror neurons. Utilizing a duration judgment task, Yamamoto and Miura (2012) provide a positive answer that stationary images potentially dilate subjective time when they imply the running rather than the standing of a character. Such time effect suggests enhanced responses of mirror neurons induced by a static image with implied biological motion. In another study with point-light walker animation (Wang and Jiang, 2012), time expansion is associated with stimuli containing biological motion components, even when observers are unaware of their biological nature. Thus, apparent duration serves as a sensitive assessment of life-relevant signals and their neural substrates. However, one has to be cautious when employing subjective time as a tool. Various factors including stimulus features, contexts, and drug administration are known to influence time perception (Meck, 1996; Eagleman, 2008) and distributed brain areas are suggested to engage in the

coding of subjective time (Mauk and Buonomano, 2004; Merchant et al., 2013). Therefore, it is a challenge to specify neural activities which are reflected in perceived time. This limitation calls for careful design of experiments to control possible confounding factors. Consulting well-established physiological and functional properties of certain neural structures may provide help in using perceived time as a research tool.

There is another advantage to measuring apparent time in psychological and neuroscience research. Object features and identities can be processed rapidly within the first tens to hundreds milliseconds after stimulus onsets (Thorpe et al., 1996). Thus, conventional RTs primarily reflect the state of transient neural responses caused by onsets. On the other hand, duration perception needs more information that can track the lapse of event. Not only onset transients, but also sustained responses and offset transients are involved in the temporal computation. Neurophysiological studies have argued that gradual changes of neural responses, e.g., ramping activities, over frontal and parietal areas serve the neural representation of mental time (see a review by Wittmann, 2013). Although such notion needs further elaboration to accommodate current doubts and inconsistencies, it is clear that certain sustained neural responses are forwarded as time signals which could be used for phenomenal representation. To this end, apparent time serves as a tool to assess, at least partially, the sustained neural responses between onset and offset transients. However, it is important to notice that there is currently no clear delineation between the contributions of transient and sustained signals to subjective time. Considering that properties of onset transients, mainly their amplitudes, profoundly alter the representation of brief time intervals (Noguchi and Kakigi, 2006; Terao

et al., 2008), this challenge will significantly limit the application of subjective time in studies of sustained neural responses.

## Concluding Remarks

Perceptual organization is an essential component of the efficient coding of the world. Here, we highlight its role in modulating the temporal representation of objects and explore its biological consequences which might underlie experiences of brief time. As a way to structure the sensory inputs, perceptual organization modifies the neural responses as well as the homeostatic conditions induced by an object. Altered biological states may change the information that is transformed into a feeling of lasting long or short (Eagleman and Pariyadath, 2009; Wittmann, 2009; Zhou et al., 2014a). In this view, perceptual organization unifies a number of time illusions on a common principle and further enables us to predict the apparent duration in a given context. The retrospective and context-dependent nature also suggests that the link between perceptual organization and subjective time is applicable in time judgment not only for online objects but also for memorized events. Using subjective time as a research tool, on the other hand, complements other behavioral measurements such as RT paradigms. With this method, it is also possible to investigate both transient and sustained neural responses encoding object identity.

## Acknowledgments

The work is supported by the National Natural Science Foundation of China (Projects 31100735) and the National Basic Research Program of China (973 Program 2015CB351800).

## References

- Aaen-Stockdale, C., Hotchkiss, J., Heron, J., and Whitaker, D. (2011). Perceived time is spatial frequency dependent. *Vision Res.* 51, 1232–1238. doi: 10.1016/j.visres.2011.03.019
- Alards-Tomalin, D., Leboe-McGowan, J. P., Shaw, J. D. M., and Leboe-McGowan, L. C. (2014). The effects of numerical magnitude, size, and color saturation on perceived interval duration. *J. Exp. Psychol. Learn. Mem. Cogn.* 40, 555–566. doi: 10.1037/a0035031
- Alexander, I., Thilo, K. V., Cowey, A., and Walsh, V. (2005). Chronostasis without voluntary action. *Exp. Brain Res.* 161, 125–132. doi: 10.1007/s00221-004-2054-3
- Bao, Y., Szymaszek, A., Wang, X., Oron, A., Pöppel, E., and Szelag, E. (2013). Temporal order perception of auditory stimuli is selectively modified by tonal and non-tonal language environments. *Cognition* 129, 579–585. doi: 10.1016/j.cognition.2013.08.019
- Bruno, A., Ayhan, I., and Johnston, A. (2010). Retinotopic adaptation-based visual duration compression. *J. Vis.* 10, 30. doi: 10.1167/10.10.30
- Burr, D., Tozzi, A., and Morrone, M. C. (2007). Neural mechanisms for timing visual events are spatially selective in real-world coordinates. *Nat. Neurosci.* 10, 423–425. doi: 10.1038/nn1874
- de Haan, B., and Rorden, C. (2010). Similarity grouping and repetition blindness are both influenced by attention. *Front. Hum. Neurosci.* 4:20. doi: 10.3389/fnhum.2010.00020
- Donner, K., and Fagerholm, P. (2003). Visual reaction time: neural conditions for the equivalence of stimulus area and contrast. *Vision Res.* 43, 2937–2940. doi: 10.1016/S0042-6989(03)00472-3
- Eagleman, D. M. (2008). Human time perception and its illusions. *Curr. Opin. Neurobiol.* 18, 131–136. doi: 10.1016/j.conb.2008.06.002
- Eagleman, D. M., and Pariyadath, V. (2009). Is subjective duration a signature of coding efficiency? *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 1841–1851. doi: 10.1098/rstb.2009.0026
- Fraisse, P. (1984). Perception and estimation of time. *Annu. Rev. Psychol.* 35, 1–36. doi: 10.1146/annurev.ps.35.020184.000245
- Geiser, E., and Gabrieli, J. D. E. (2013). Influence of rhythmic grouping on duration perception: a novel auditory illusion. *PLoS ONE* 8:e54273. doi: 10.1371/journal.pone.0054273
- Gibson, J. J. (1975). “Events are perceivable but time is not,” in *The Study of Time II*, eds J. T. Fraser and N. Lawrence (Berlin: Springer), 295–301.
- Grice, G. R., Nullmeyer, R., and Spiker, V. A. (1982). Human reaction time: toward a general theory. *J. Exp. Psychol. Gen.* 111, 135–153. doi: 10.1037/0096-3445.111.1.135
- Han, S., Jiang, Y., Mao, L., Humphreys, G. W., and Gu, H. (2005). Attentional modulation of perceptual grouping in human visual cortex: functional MRI studies. *Hum. Brain Mapp.* 25, 424–432. doi: 10.1002/hbm.20119
- Herzog, M. H., and Fahle, M. (2002). Effects of grouping in contextual modulation. *Nature* 415, 433–436. doi: 10.1038/415433a
- Hodinott-Hill, I., Thilo, K. V., Cowey, A., and Walsh, V. (2002). Auditory chronostasis: hanging on the telephone. *Curr. Biol.* 12, 1779–1781. doi: 10.1016/S0960-9822(02)01219-8
- Hunt, A. R., Chapman, C. S., and Kingstone, A. (2008). Taking a long look at action and time perception. *J. Exp. Psychol. Hum. Percept. Perform.* 34, 125–136. doi: 10.1037/0096-1523.34.1.125
- Jazayeri, M., and Shadlen, M. N. (2010). Temporal context calibrates interval timing. *Nat. Neurosci.* 13, 1020–1026. doi: 10.1038/nn.2590
- Johnston, A., Arnold, D. H., and Nishida, S. (2006). Spatially localized distortions of event time. *Curr. Biol.* 16, 472–479. doi: 10.1016/j.cub.2006.01.032
- Kimchi, R., and Hadad, B.-S. (2002). Influence of past experience on perceptual grouping. *Psychol. Sci.* 13, 41–47. doi: 10.1111/1467-9280.00407
- Kramer, P., and Yantis, S. (1997). Perceptual grouping in space and time: evidence from the Ternus display. *Percept. Psychophys.* 59, 87–99. doi: 10.3758/BF03206851

- MacDonald, C. J., and Meck, W. H. (2004). Systems-level integration of interval timing and reaction time. *Neurosci. Biobehav. Rev.* 28, 747–769. doi: 10.1016/j.neubiorev.2004.09.007
- Matthews, W. J. (2011a). How do changes in speed affect the perception of duration? *J. Exp. Psychol. Hum. Percept. Perform.* 37, 1617–1627. doi: 10.1037/a0022193
- Matthews, W. J. (2011b). Stimulus repetition and the perception of time: the effects of prior exposure on temporal discrimination, judgment, and production. *PLoS ONE* 6:e19815. doi: 10.1371/journal.pone.0019815
- Matthews, W. J., Stewart, N., and Wearden, J. H. (2011). Stimulus intensity and the perception of duration. *J. Exp. Psychol. Hum. Percept. Perform.* 37, 303–313. doi: 10.1037/a0019961
- Mauk, M. D., and Buonomano, D. V. (2004). The neural basis of temporal processing. *Annu. Rev. Neurosci.* 27, 307–340. doi: 10.1146/annurev.neuro.27.070203.144247
- Meck, W. H. (1996). Neuropharmacology of timing and time perception. *Cogn. Brain Res.* 3, 227–242. doi: 10.1016/0926-6410(96)00009-2
- Merchant, H., Harrington, D. L., and Meck, W. H. (2013). Neural basis of the perception and estimation of time. *Annu. Rev. Neurosci.* 36, 313–336. doi: 10.1146/annurev-neuro-062012-170349
- New, J. J., and Scholl, B. J. (2009). Subjective time dilation: spatially local, object-based, or a global visual experience? *J. Vis.* 9, 4, 1–11. doi: 10.1167/9.2.4
- Nisly, S. J., and Wasserman, G. S. (1989). Intensity dependence of perceived duration: data, theories, and neural integration. *Psychol. Bull.* 106, 483–496. doi: 10.1037/0033-2909.106.3.483
- Noguchi, Y., and Kakigi, R. (2006). Time representations can be made from non-temporal information in the brain: an MEG study. *Cereb. Cortex* 16, 1797–1808. doi: 10.1093/cercor/bhj117
- Ohyama, J., and Watanabe, K. (2007). Unpredictable visual changes cause temporal memory averaging. *Vision Res.* 47, 2727–2731. doi: 10.1016/j.visres.2007.06.020
- Ono, F., and Kawahara, J.-I. (2007). The subjective size of visual stimuli affects the perceived duration of their presentation. *Percept. Psychophys.* 69, 952–957. doi: 10.3758/BF03193932
- Ono, F., and Kitazawa, S. (2009). The effect of marker size on the perception of an empty interval. *Psychon. Bull. Rev.* 16, 182–189. doi: 10.3758/PBR.16.1.182
- Orgs, G., Bestmann, S., Schuur, F., and Haggard, P. (2011). From body form to biological motion: the apparent velocity of human movement biases subjective time. *Psychol. Sci.* 22, 712–717. doi: 10.1177/0956797611406446
- Palmer, J., Huk, A. C., and Shadlen, M. N. (2005). The effect of stimulus strength on the speed and accuracy of a perceptual decision. *J. Vis.* 5, 376–404. doi: 10.1167/5.5.1
- Pariyadath, V., and Eagleman, D. (2007). The effect of predictability on subjective duration. *PLoS ONE* 2:e1264. doi: 10.1371/journal.pone.0001264
- Pariyadath, V., and Eagleman, D. M. (2012). Subjective duration distortions mirror neural repetition suppression. *PLoS ONE* 7:e49362. doi: 10.1371/journal.pone.0049362
- Park, J., Schlag-Rey, M., and Schlag, J. (2003). Voluntary action expands perceived duration of its sensory consequence. *Exp. Brain Res.* 149, 527–529. doi: 10.1007/s00221-003-1376-x
- Peterson, D. J., and Berryhill, M. E. (2013). The Gestalt principle of similarity benefits visual working memory. *Psychon. Bull. Rev.* 20, 1282–1289. doi: 10.3758/s13423-013-0460-x
- Pöppel, E. (1989). “Taxonomy of the subjective: an evolutionary perspective,” in *Neuropsychology of Visual Perception*, ed J. W. Brown (Hillsdale NJ: Lawrence Erlbaum Associates), 219–232.
- Pöppel, E. (1997). A hierarchical model of temporal perception. *Trends Cogn. Sci.* 1, 56–61. doi: 10.1016/S1364-6613(97)01008-5
- Pöppel, E. (2009). Pre-semantically defined temporal windows for cognitive processing. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 1887–1896. doi: 10.1098/rstb.2009.0015
- Posner, M. I. (2005). Timing the brain: mental chronometry as a tool in neuroscience. *PLoS Biol.* 3:e51. doi: 10.1371/journal.pbio.0030051
- Rizzolatti, G., and Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.* 27, 169–192. doi: 10.1146/annurev.neuro.27.070203.144230
- Rose, D., and Summors, J. (1995). Duration illusions in a train of visual stimuli. *Perception* 24, 1177–1187. doi: 10.1068/p241177
- Sasaki, K., Yamamoto, K., and Miura, K. (2013). The difference in speed sequence influences perceived duration. *Perception* 42, 198–207. doi: 10.1068/p7241
- Schindel, R., Rowlands, J., and Arnold, D. H. (2011). The oddball effect: perceived duration and predictive coding. *J. Vis.* 11, 17. doi: 10.1167/11.2.17
- Terao, M., Watanabe, J., Yagi, A., and Nishida, S. (2008). Reduction of stimulus visibility compresses apparent time intervals. *Nat. Neurosci.* 11, 541–542. doi: 10.1038/nn.2111
- Thorpe, L. A., and Trehub, S. E. (1989). Duration illusion and auditory grouping in infancy. *Dev. Psychol.* 25, 122–127. doi: 10.1037/0012-1649.25.1.122
- Thorpe, S., Fize, D., and Marlot, C. (1996). Speed of processing in the human visual system. *Nature* 381, 520–522. doi: 10.1038/381520a0
- Tse, P. U., Intriligator, J., Rivest, J., and Cavanagh, P. (2004). Attention and the subjective expansion of time. *Percept. Psychophys.* 66, 1171–1189. doi: 10.3758/BF03196844
- Usher, M., and McClelland, J. L. (2001). The time course of perceptual choice: the leaky, competing accumulator model. *Psychol. Rev.* 108, 550–592. doi: 10.1037/0033-295X.108.3.550
- van Wassenhove, V., Buonomano, D. V., Shimojo, S., and Shams, L. (2008). Distortions of subjective time perception within and across senses. *PLoS ONE* 3:e1437. doi: 10.1371/journal.pone.0001437
- Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., et al. (2012). A century of Gestalt psychology in visual perception I. Perceptual grouping and figure-ground organization. *Psychol. Bull.* 138, 1172–1217. doi: 10.1037/a0029333
- Wang, L., and Jiang, Y. (2012). Life motion signals lengthen perceived temporal duration. *Proc. Natl. Acad. Sci. U.S.A.* 109, E673–E677. doi: 10.1073/pnas.1115515109
- Wittmann, M. (2009). The inner experience of time. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 1955–1967. doi: 10.1098/rstb.2009.0003
- Wittmann, M. (2013). The inner sense of time: how the brain creates a representation of duration. *Nat. Rev. Neurosci.* 14, 217–223. doi: 10.1038/nrn3452
- Wittmann, M., van Wassenhove, V., Craig, A. D., and Paulus, M. P. (2010). The neural substrates of subjective time dilation. *Front. Hum. Neurosci.* 4:2. doi: 10.3389/neuro.09.002.2010
- Woodrow, H. (1951). “Time perception,” in *Handbook of Experimental Psychology*, ed S. S. Stevens (Oxford: Wiley), 1224–1236.
- Xuan, B., Zhang, D., He, S., and Chen, X. (2007). Larger stimuli are judged to last longer. *J. Vis.* 7, 1–5. doi: 10.1167/7.10.2
- Yamamoto, K., and Miura, K. (2012). Time dilation caused by static images with implied motion. *Exp. Brain Res.* 223, 311–319. doi: 10.1007/s00221-012-3259-5
- Yarrow, K., Haggard, P., Heal, R., Brown, P., and Rothwell, J. C. (2001). Illusory perceptions of space and time preserve cross-saccadic perceptual continuity. *Nature* 414, 302–305. doi: 10.1038/35104551
- Yarrow, K., Johnson, H., Haggard, P., and Rothwell, J. C. (2004). Consistent chronostasis effects across saccade categories imply a subcortical efferent trigger. *J. Cogn. Neurosci.* 16, 839–847. doi: 10.1162/089982904970780
- Yarrow, K., and Rothwell, J. C. (2003). Manual chronostasis: tactile perception precedes physical contact. *Curr. Biol.* 13, 1134–1139. doi: 10.1016/S0960-9822(03)00413-5
- Zemel, R. S., Behrmann, M., Mozer, M. C., and Bavelier, D. (2002). Experience-dependent perceptual grouping and object-based attention. *J. Exp. Psychol. Hum. Percept. Perform.* 28, 202–217. doi: 10.1037/0096-1523.28.1.202
- Zhou, B., Pöppel, E., and Bao, Y. (2014a). In the jungle of time: the concept of identity as a way out. *Front. Psychol.* 5:844. doi: 10.3389/fpsyg.2014.00844
- Zhou, B., Qin, J., Mao, L., Han, S., and Pöppel, E. (2010). Modulations of temporal perception by consciously and unconsciously perceived stimuli. *Perception* 39, 900–908. doi: 10.1068/p6662
- Zhou, B., Yang, S., Mao, L., and Han, S. (2014b). Visual feature processing in the early visual cortex affects duration perception. *J. Exp. Psychol. Gen.* 143, 1893–1902. doi: 10.1037/a0037294

**Conflict of Interest Statement:** 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.

Copyright © 2015 Zhou, Zhang and Mao. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

