Cognition 171 (2018) 52–64

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Original Articles

Expected but omitted stimuli affect crossmodal interaction

Marcello Costantini[a,b,c,*], Daniele Migliorati[b,c], Brunella Donno[b,c], Miroslav Sirota[a], Francesca Ferri[a,*]

[a] Centre for Brain Science, Department of Psychology, University of Essex, Colchester, UK
[b] Laboratory of Neuropsychology and Cognitive Neuroscience, Department of Neuroscience and Imaging, University G. d'Annunzio, Chieti, Italy
[c] Institute for Advanced Biomedical Technologies – ITAB, Foundation University G. d'Annunzio, Chieti, Italy

A R T I C L E I N F O

Keywords:
Crossmodal congruency effect
Predictive coding
Sensory expectation

A B S T R A C T

One of the most important ability of our brain is to integrate input from different sensory modalities to create a coherent representation of the environment. Does expectation affect such multisensory integration? In this paper, we tackled this issue by taking advantage from the crossmodal congruency effect (CCE). Participants made elevation judgments to visual target while ignoring tactile distractors. We manipulated the expectation of the tactile distractor by pairing the tactile stimulus to the index finger with a high-frequency tone and the tactile stimulus to the thumb with a low-frequency tone in 80% of the trials. In the remaining trials we delivered the tone and the visual target, but the tactile distractor was omitted (Study 1). Results fully replicated the basic crossmodal congruency effect. Strikingly, the CCE was observed, though at a lesser degree, also when the tactile distractor was not presented but merely expected. The contingencies between tones and tactile distractors were reversed in a follow-up study (Study 2), and the effect was further tested in two conceptual replications using different combinations of stimuli (Studies 5 and 6). Two control studies ruled out alternative explanations of the observed effect that would not involve a role for tactile distractors (Studies 3, 4). Two additional control studies unequivocally proved the dependency of the CCE on the spatial and temporal expectation of the distractors (Study 7, 8). An internal small-scale meta-analysis showed that the crossmodal congruency effect with predicted distractors is a robust medium size effect. Our findings reveal that multisensory integration, one of the most basic and ubiquitous mechanisms to encode external events, benefits from expectation of sensory input.

1. Introduction

Two key mechanisms help us to cope with an overwhelming amount of sensory inputs coming from the environment: sensory expectation and crossmodal interaction. The former refers to the idea that we do not solely react to external stimuli; rather we constantly create predictions about forthcoming sensory events (Engel, Fries, & Singer, 2001). The latter refers to the idea that we do not use sensory systems one at a time, rather we simultaneously process information coming from different sensory modalities. These two mechanisms can be observed already in non-human primates (Amemori & Sawaguchi, 2006; Siemann et al., 2014), suggesting that they are unlikely related to the privileged cognitive status of humans. Conversely, they might represent a fundamental prerequisite for an efficient interaction with the environment.

Models of predictive brain have been used to explain how expectation of upcoming stimuli is generated (for reviews: (Clark, 2013; Friston, 2010). According to these models, expectation at the neural level takes the form of increased baseline neural activity (i.e., biased by the probability of stimulus occurrence) and increased evoked response (i.e., similar for expected and actual stimuli; for a review see: (Summerfield & de Lange, 2014). For instance, previous research has shown that cues predicting a forthcoming visual stimulus lead to increases in BOLD signal in category-specific visual regions. For example, when the word 'house' predicts the subsequent occurrence of a house, it triggers higher BOLD signals in the parahippocampal place area (Puri, Wojciulik, & Ranganath, 2009). Behavioral evidence demonstrates that expectation is beneficial for processing and responding to external stimuli. For instance, expectation of low-level features (e.g., colour, direction of motion) leads to facilitated processing of stimuli containing those features (Ball & Sekuler, 1981; Corbetta, Miezin, Dobmeyer, Shulman, & Petersen, 1990; Saenz, Buracas, & Boynton, 2002).

Only recent research has begun to investigate the relationship between expectation and crossmodal interaction. Examining this issue is critical for the understanding of how we perceive and react to environmental stimuli. Indeed, in daily life we usually do not perceive external events through only one sensory modality. Instead,

* Corresponding authors at: Centre for Brain Science, Department of Psychology, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, UK.
E-mail addresses: mcostaa@essex.ac.uk (M. Costantini), fferri@essex.ac.uk (F. Ferri).

http://dx.doi.org/10.1016/j.cognition.2017.10.016
Received 16 September 2016; Received in revised form 19 October 2017; Accepted 20 October 2017
0010-0277/

information from the environment reaches us via multiple sensory systems. Integrating information across the senses improves a wide range of behavioral outcomes, including detection (Lovelace, Stein, & Wallace, 2003; Stein & Wallace, 1996), localization (Nelson et al., 1998; Wilkinson, Meredith, & Stein, 1996), and speed of response (Diederich & Colonius, 2004; Hershenson, 1962).

To date, however, research on the relationship between expectation and crossmodal interaction has focused predominantly on the extent to which top-down expectations impact on actual multimodal events (Gau & Noppeney, 2016; Nahorna, Berthommier, & Schwartz, 2012) as in the case of the McGurk effect. In the McGurk effect, participants are presented with the auditory phoneme /ga/ synchronous with an incongruent lip movement /ba/. This leads to the illusory perception of a different syllable /da/ (McGurk & MacDonald, 1976). Nahorna et al. (2012) manipulated participants' top-down congruency expectations by presenting McGurk stimuli embedded in blocks of congruent or incongruent syllables. They showed that the multimodal McGurk effect was largely reduced when the constituent unisensory stimuli were preceded by an incoherent audiovisual context. Other studies (Stekelenburg & Vroomen, 2012) have shown similar top-down effects on crossmodal interaction employing different pair of stimuli, e.g. audio-tactile or visuo-tactile. For instance, by using a spatial cuing task, Spence and Driver (Spence & Driver, 1996) showed that participants were faster at judging the elevation of visual or auditory targets when the location of the upcoming stimulus was cued by a stimulus in either the same or different sensory modality. This finding is intriguing and clearly suggests that expectation of sensory events might occur across modalities.

So far, little research has investigated whether crossmodal interaction between an actual stimulus and an expected, but omitted stimulus, could occur. In support of this working hypothesis, there is evidence showing that expectations affect the sensory response in the absence of sensory input (den Ouden, Friston, Daw, McIntosh, & Stephan, 2009; Kok, Rahnev, Jehee, Lau, & de Lange, 2012; SanMiguel, Widmann, Bendixen, Trujillo-Barreto, & Schroger, 2013; Todorovic, van Ede, Maris, & de Lange, 2011; Wacongne et al., 2011). For instance, Kok, Failing, and de Lange (2014) showed that expectation of a specific visual stimulus evokes a pattern of activity in the visual cortex with similar features as those evoked by real stimuli.

Starting from this evidence, we seek to investigate whether expectation of a tactile event is a sufficient condition to elicit crossmodal interaction in a modified version of the crossmodal distractor congruency task (Driver & Spence, 1998; Spence, 2010, 2011; Spence, Pavani, & Driver, 2000, 2004; Spence & Walton, 2005). In a typical study, participants hold two foam blocks, one in either hand, provided with vibrotactile stimulators and light emitting diodes (LEDs) in the upper and lower surfaces. On each trial, a vibrotactile and a visual stimulus are presented randomly from any one of the four possible stimulus locations. Participants are required to make speeded elevation (up/down) discriminations for each vibrotactile target stimulus, presented to either the index finger or the thumb, while simultaneously ignoring any visual distractor. The common finding is that participants are significantly faster at discriminating the elevation of tactile targets when visual distractors are presented at congruent elevation. That is, a spatially non-predictive visual cue enhances judgments for tactile targets presented near to it, relative to those presented elsewhere (Driver & Spence, 1998). The effect is consistently found also when participants respond to the visual stimulus trying to ignore the tactile stimulus (Spence & Walton, 2005).

In this study we conceived a modified version of the classic crossmodal distractor congruency task. Participants made elevation judgments to visual target while ignoring tactile distractors. We manipulated the expectation of the tactile distractor by pairing the tactile stimulus to the index finger with a high-frequency tone and the tactile stimulus to the thumb with a low-frequency tone in 80% of the trials. In the remaining trials we delivered the tone and the visual target, but the tactile distractor was omitted.

Based on evidence suggesting that expected stimuli evoke a pattern of activity with similar features as that evoked by the real stimulus (Kok et al., 2014), we predict a crossmodal congruency effect not only when the *actual tactile distractor* is spatially incongruent with the visual target, but also when the *expected tactile distractor* is spatially incongruent with the visual target, even if omitted (study 1).

The contingencies between tones and tactile distractors were reversed in a follow-up study (study 2). Two additional control studies (studies 3 & 4) aimed at testing the dependency of our effect on the expectation of the tactile distractor. These studies ruled out the possibility that the mere association between the auditory cue and visual target could account for by our results (study 3) by investigating the time course of the audio-visual crossmodal congruency effect (study 4). Two conceptual replications of study 1 (studies 5 & 6) aimed at assessing the generalizability of our effect to other stimulus combinations. In particular, in study 5 the cue, the distractor and the target were auditory, visual and tactile, respectively; while in study 6 the cue, the distractor and the target were visual, visual and tactile, respectively. Two further control studies (studies 7 & 8) aimed at testing the dependency of our effect on the spatial and temporal predictability of the distractor. In study 7 the cue, the distractor and the target were auditory, visual and tactile, respectively, as in study 5, but in this case the auditory cue predicted the spatial location of the forthcoming distractor with a 50% of accuracy, thus making it “spatially” unpredictable. Finally, in study 8 the cue, the distractor and the target were auditory, visual and tactile, respectively. In this study, the expectation of the distractor dissipated over time in specific trials, due to the delayed presentation of the target. Thus, the crossmodal congruency effect should not be observed in these trials.

2. Methodological disclosure and description of the analysis

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in each study. We report all the studies conducted for this project (Simmons, Nelson, & Simonsohn, 2011). We have used two inferential frameworks to assess the evidence for the critical effects: null hypothesis significance testing and a Bayesian inference framework – Bayes factor analysis. The latter framework enabled us to quantified relative evidence to support the null effect model against models assuming an effect (or vice versa). The crucial notion here is a Bayes factor (BF), which is the ratio of the probability of the data given model A (e.g., the null model) to the probability of the data given model B (e.g., a model assuming a certain distribution of effects). Bayes factors allows us to quantify how many more times are the data likely to occur under the assumption of the model A compared to the assumptions of model B (or vice versa). For instance, $BF_{01} = 20$ means that the data are 20 times more likely to occur under the model A (i.e., the null model here) relative to the model B. A Bayes Factor, BF_{01} with a value lower than 1 indicates that the model assuming the effect (model B) is more likely relative to the null model (model A) and with value greater than 1 indicates that the model assuming no effect is more likely relative to the model assuming the effect. Furthermore, the Bayes Factor values may also be interpreted as evidence categories, for example, BF_{01} values between 1 and 3 indicate *anecdotal evidence* to support the null model relative to the competing model, whereas values greater than 100 indicate *extreme evidence* to support the null model (Jeffreys, 1961; Lee & Wagenmakers, 2014). Finally, if the prior probability odds are defined, then Bayes factors can be combined into the posterior odds and can thus quantify support for tested hypotheses. For instance, if we assume prior odds of the two competing models to be 1 (i.e., equally likely) before running a study, then $BF_{01} = 100$ can be combined into the posterior odds ($1 * 100 = 100$), which will mean that the null model is 100 more likely relative to the compared model (i.e., assuming a specific distribution of the effects). Here, we calculated a default Bayes factors using JASP and R package Bayes Factor (Love