Cognition 170 (2018) 130–146

Contents lists available at ScienceDirect

Cognition

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

Original Articles

The role of multisensory interplay in enabling temporal expectations

Felix Ball[a,b,c,*], Lara E. Michels[a], Carsten Thiele[a], Toemme Noesselt[a,c]

[a] Department of Biological Psychology, Otto-von-Guericke University, Magdeburg, Germany
[b] Department of Neurology, Otto-von-Guericke University, Magdeburg, Germany
[c] Center of Behavioural Brain Sciences, Otto-von-Guericke University, Magdeburg, Germany

ARTICLE INFO

Abbreviations:
TE
temporal expectation/temporal expectancy
IE
inverse efficiency
MSI
multisensory interplay
EEG
electroencephalography
A
audio/auditory
V
visual
AV
audiovisual
RT
response time

Keywords:
Temporal expectation
Temporal orienting
Multisensory interplay
Redundant target
Spatial coincidence
Auditory dominance

ABSTRACT

Temporal regularities can guide our attention to focus on a particular moment in time and to be especially vigilant just then. Previous research provided evidence for the influence of temporal expectation on perceptual processing in unisensory auditory, visual, and tactile contexts. However, in real life we are often exposed to a complex and continuous stream of multisensory events. Here we tested – in a series of experiments – whether temporal expectations can enhance perception in multisensory contexts and whether this enhancement differs from enhancements in unisensory contexts. Our discrimination paradigm contained near-threshold targets (subject-specific 75% discrimination accuracy) embedded in a sequence of distractors. The likelihood of target occurrence (early or late) was manipulated block-wise. Furthermore, we tested whether spatial and modality-specific target uncertainty (i.e. predictable vs. unpredictable target position or modality) would affect temporal expectation (TE) measured with perceptual sensitivity (d′) and response times (RT). In all our experiments, hidden temporal regularities improved performance for expected multisensory targets. Moreover, multisensory performance was unaffected by spatial and modality-specific uncertainty, whereas unisensory TE effects on d′ but not RT were modulated by spatial and modality-specific uncertainty. Additionally, the size of the temporal expectation effect, i.e. the increase in perceptual sensitivity and decrease of RT, scaled linearly with the likelihood of expected targets. Finally, temporal expectation effects were unaffected by varying target position within the stream. Together, our results strongly suggest that participants quickly adapt to novel temporal contexts, that they benefit from multisensory (relative to unisensory) stimulation and that multisensory benefits are maximal if the stimulus-driven uncertainty is highest. We propose that enhanced informational content (i.e. multisensory stimulation) enables the robust extraction of temporal regularities which in turn boost (uni-)sensory representations.

1. Introduction

The amount of information organisms are confronted with at any given moment is tremendous. It is therefore imperative to focus on particular aspects of the incoming information and to preferentially process the most relevant parts as both information overflow and missing important bits of information can have severe consequences (e.g. in traffic). Spatial attention offers one solution to selectively increase the salience of particular information and has been the focus of numerous previous investigations (Ball, Bernasconi, & Busch, 2015; Ball & Busch, 2015; Ball, Elzemann, & Busch, 2014; Kovalenko & Busch, 2016; Luck et al., 2004; Posner, 1980; Posner, Snyder, & Davidson, 1980; Yeshurun & Carrasco, 1998). Another way to facilitate information processing is to anticipate when future objects and events may occur and what these events/objects might be (the *what*: Baylis & Driver, 1993; Behrmann, Zemel, & Mozer, 1998; Chen, 2000; Duncan, 1984; Kramer, Weber, & Watson, 1997; Vecera & Farah, 1994 and *when*: Correa, Lupiáñez, Milliken, & Tudela, 2004; Correa, Lupiáñez, Madrid, & Tudela, 2006; Coull & Nobre, 2008; Doherty, Rao, Mesulam, & Nobre, 2005; Nobre, 2001; Rohenkohl, Coull, & Nobre, 2011; Rohenkohl, Gould, Pessoa, & Nobre, 2014). In this article we will differentiate between different aspects of temporal information influencing behaviour. The term temporal predictability will be used to denote exogenous factors, e.g. the manipulation of temporal regularities by experimental design. Endogenous factors derived from these objective temporal regularities – i.e. temporal expectations (TE) generated by the participant – will be referred to as temporal attention or temporal expectation (in accord with e.g. Bendixen,

* Corresponding author at: Otto-von-Guericke-University, Institute of Psychology, Department of Biological Psychology, Universitätsplatz 2, 39106 Magdeburg, Germany
E-mail address: felix.ball@ovgu.de (F. Ball).

http://dx.doi.org/10.1016/j.cognition.2017.09.015
Received 21 October 2016; Received in revised form 22 September 2017; Accepted 26 September 2017

SanMiguel, & Schröger, 2012).

Previous research on temporal attention preferentially used three main paradigms (see Nobre & Rohenkohl, 2014, for a recent review) which have been based on rhythmic variations, temporal cueing, and foreperiod duration. In studies using rhythmic variations, temporal expectations are automatically generated by presenting an isochronous stimulus sequence (Cravo, Rohenkohl, Wyart, & Nobre, 2013; Jones, Moynihan, MacKenzie, & Puente, 2002; Mathewson, Fabiani, Gratton, Beck, & Lleras, 2010; Rohenkohl, Cravo, Wyart, & Nobre, 2012; Sanabria, Capizzi, & Correa, 2011). Target stimuli are either shown at the end of or are embedded within the rhythmic sequence. Only targets presented in phase with the rhythm are temporally predictable, while arrhythmically presented targets are unpredictable. In temporal cueing experiments (Correa et al., 2004; Coull & Nobre, 1998; Griffin, Miniussi, & Nobre, 2001, 2002; Jepma, Wagenmakers, & Nieuwenhuis, 2012; Miniussi, Wilding, Coull, & Nobre, 1999) a signal predicts the delay between cue and target (e.g. 200 ms vs. 800 ms) with a certain probability (e.g. 75%), in close resemblance to spatial cueing paradigms (Posner, 1980; Posner et al., 1980). Here, TE can be manipulated on a trial-by-trial basis, whereas belief about cue validity builds up over time. There is corroborating evidence from both rhythm and cueing studies indicating that temporal predictability of events enables us to create temporal expectations which in turn improve performance: they enhance detectability of targets, increase accuracy in discrimination tasks (e.g. frequency judgement), and decrease response times (Nobre & Rohenkohl, 2014). The third approach investigating TE utilizes foreperiod paradigms (Lange & Röder, 2006; Lange, Rösler, & Röder, 2003; Niemi & Näätänen, 1981; Rolke & Hofmann, 2007; Westheimer & Ley, 1996) in which hazard rates – the conditional probability of the occurrence of a target given that it has not yet been presented – are manipulated (Nobre & Rohenkohl, 2014). In particular, the cue-target delay (i.e. the foreperiod) is varied between blocks (e.g. short or long foreperiod); temporal regularities are not explicitly cued, thus temporal expectation builds up over trials. In these studies performance consistently decreases with increasing foreperiod duration, and it has been suggested that this might be due to participant's decreased temporal precision or participant's higher temporal uncertainty with increasing cue-target intervals (Klemmer, 1956; Näätänen, Muranen, & Merisalo, 1974; Näätänen & Merisalo, 1977; Niemi & Näätänen, 1981).

The paradigms above all have in common that the effects of temporal attention were tested implicitly – i.e. knowledge about time-of-target-occurrence was not explicitly assessed – but nevertheless, the temporal predictable context improved performance. Another line of research directly investigated the representation of time using temporal bisection tasks and switch paradigms (Akdoğan & Balcı, 2016; Balci, Freestone, & Gallistel, 2009; Balcı et al., 2011; Bogacz, Brown, Moehlis, Holmes, & Cohen, 2006; Çavdaroğlu, Zeki, & Balcı, 2014; Çoşkun, Sayalı, Gürbüz, & Balcı, 2015; Freestone, Balcı, Simen, & Church, 2015) among other tasks. Results from both human and animal studies revealed that participants were able to base their temporal decisions on – sometimes noisy – time estimates. The noise intrinsic in these time estimates can be due to exogenous factors (variability of external sources) and additionally due to the endogenous properties of the temporal representations. Concordantly, several computational models have been put forward to account for the observed effects including pacemaker accumulators and drift diffusion models (see e.g. for a recent review Balcı & Simen, 2016). Given several similarities between explicit and implicit timing results, intrinsic temporal estimators such as pacemaker accumulators might be used for both, the explicit and implicit use of temporal regularities.

Another similarity of the paradigms mentioned above is that they investigate temporal attention explicitly or implicitly but in the absence of additional potentially distracting information. Indeed, in most of these studies, the target is presented in isolation and can easily be perceived as target (e.g. targets are colour coded, presented at the end of sequences, or presented in isolation after the cue, and thus are quite obvious). In the last years, novel paradigms have been designed to create more ecologically valid contexts with distracting information and with targets which are less obvious (e.g. Jaramillo & Zador, 2011; Shen & Alain, 2011). Among them are attentional blink studies (stimulus sequences with an embedded target and probe; e.g. Shen & Alain, 2011, 2012) and studies combining foreperiod with rhythmic designs in which the hazard rate of targets – which themselves are hidden in a sequence of distracting stimuli – varies (Jaramillo & Zador, 2011).

A different promising approach to investigate temporal expectation in more ecologically valid contexts could include the use of multisensory stimuli, as many real-life events stimulate more than one sense. Concordantly, there is evidence that seeing lip movements can enhance speech perception (Grant & Greenberg, 2001; Reisberg, Mclean, & Goldfield, 1987; Risberg & Lubker, 1978; Sumby & Pollack, 1954) and that multisensory perception also improves later memory retrieval (Luria, 1968; Shams & Seitz, 2008). Moreover, several psychophysical studies indicate that redundant multisensory stimulation can improve performance relative to unisensory stimulation (Alais & Burr, 2004; Driver & Noesselt, 2008; Forster, Cavina-Pratesi, Aglioti, & Berlucchi, 2002; Gondan, Niederhaus, Rösler, & Röder, 2005; Jaekl & Harris, 2009; Noesselt et al., 2010; Parise, Spence, & Ernst, 2012; Sinnett, Soto-Faraco, & Spence, 2008; Stevenson et al., 2014; Talsma, Doty, & Woldorff, 2007; Van der Burg, Olivers, Bronkhorst, & Theeuwes, 2008) and some have pointed at enhanced MSI with less reliable sensory input (Beauchamp, Pasalar, & Ro, 2010; Meredith & Stein, 1983; Meredith & Stein, 1986b; Werner & Noppeney, 2010) and with increasing uncertainty (Körding et al., 2007). Hence a manipulation of uncertainty or stimulus reliability should affect the strength of MSI. Concordantly, studies on visual perception modulated by sound revealed that visual sensitivity for less reliable visual stimuli is improved by simultaneously presenting an irrelevant, uninformative sound (e.g. Jaekl & Harris, 2009; Noesselt et al., 2010; Van der Burg et al., 2008), and that performance increases non-linearly when target information is doubled (presenting an audiovisual target instead of just auditory or visual target; e.g. Gondan et al., 2005). Therefore it is at least conceivable that multisensory stimulation – potentially by means of its higher informational content – can aid the statistical learning mechanisms (Barakat, Seitz, & Shams, 2013) underlying the built-up of temporal expectation. However, to our knowledge there is to date little experimental support for this hypothesis.

Several studies have looked into the relationship how spatial and modality-specific attention interacts with multisensory integration but with mixed results (e.g. Alsius, Navarra, Campbell, & Soto-Faraco, 2005; Bertelson, Vroomen, De Gelder, & Driver, 2000; Mozolic, Hugenschmidt, Peiffer, & Laurienti, 2008; Shore & Simic, 2005; Vroomen, Bertelson, & De Gelder, 2001; Werkhoven, van Erp, & Philippi, 2009). Only few studies investigated the interplay of cross-modal effects and temporal expectations (Bolger, Trost, & Schön, 2013; Jones, 2015; Lange & Röder, 2006; Menceloglu, Grabowecky, & Suzuki, 2016; Miller, Carlson, & McAuley, 2012; Mühlberg, Oriolo, & Soto-Faraco, 2014) but they focused on other aspects than the influence of multisensory stimulation on temporal expectation in their studies. For instance, Lange and Röder (2006) used a temporal attention paradigm and tested whether knowledge about temporal regularities in one modality can be transferred to another modality (though note that no combined multisensory signals were presented). In each block, participants were instructed to attend to either short or long cue-target delays and to either auditory or tactile stimuli. Lange and Röder (2006) observed shortened response times (RT) for temporally expected targets. Remarkably, they also observed that RTs were faster for stimuli in the unattended modality when presented at expected time points supporting the notion that knowledge about temporal regularities is stored as a supramodal representation (for similar findings see Bolger et al., 2013; Jones, 2015; Miller et al., 2012). Mühlberg et al. (2014) used a similar crossmodal transfer