BRILL *Multisensory Research (2017)* DOI:10.1163/22134808-00002613 brill.com/msr

Short Note

Effects of Temporal Distribution on Utility of Temporal Factors in Competitive Audio-Visual Perceived Synchrony

Jonathan M. P. Wilbiks *

Department of Psychology, Mount Allison University, 62 York Street,
Sackville, NB, E4L 1E2, Canada

Received 2 March 2017; accepted 6 September 2017

Abstract

The perception of audio-visual synchrony is affected by both temporal coincidence and stimulus congruency factors. In situations when temporal and stimulus information are not in agreement, the perceiver must rely on the relative informative value of both factors in deciding which of multiple potential binding candidates are most likely to be of a common source to a target. Previous research has shown that, all being equal, participants tend to rely primarily on temporal information, and only take stimulus information into consideration when temporal information is ambiguous. The current research seeks to examine the reliance on temporal *vs.* stimulus information by altering the degree of useful information available in temporal aspects. By varying the temporal distribution of stimuli, it was possible to either increase or decrease the number of trials on which temporal information is conclusive. Data indicate that when temporal information is less informative (i.e., when more asynchronous stimuli are presented), we become less reliant on using prior knowledge about timing relationships when making synchrony judgements. However, when temporal information is more informative (i.e., when more synchronous stimuli are presented) there is no increase in reliance on this type of information. These findings increase what is known about competitive audiovisual processing, and the fact that temporal information serves as a kind of default stimulus property, which can be decreased by reducing the utility of that information.

Keywords

Audiovisual binding, temporal distribution, stimulus effects

---

* E-mail: jwilbiks@mta.ca

DOI:10.1163/22134808-00002613

1. Introduction

Research into the perception of subjective synchrony between stimuli in different modalities has revealed a disparity in the relative weighting of visual and auditory stimuli in decision making in spatial and temporal tasks. In a spatial task, individuals are more likely to give priority to visual stimuli (i.e., visual dominance; Alais and Burr, 2004), while in a temporal task individuals are more likely to prefer auditory stimuli (i.e., auditory dominance; Burr *et al.*, 2009). This combination of reliance on different stimulus modalities for different tasks has been shown to be statistically optimal (Ernst and Banks, 2002), with additional research showing that the system abides by Bayesian logic (Ernst and Bulthoff, 2004). In previous experiments using a competitive paradigm (Wilbiks and Dyson, 2013a), two stimuli of a certain modality (e.g., vision) were presented with a 300 ms time interval between them. A single stimulus of a different modality (e.g., audition) was presented at a varying number of stimulus onset asynchronies (SOAs), ranging from coincident with the first or second of the paired stimuli, or during the time between them. The stimuli were also manipulated with regard to their stimulus congruency relationship. The main finding was that the primary factor for audio-visual perceived synchrony was temporal coincidence, with stimulus congruency only playing a role when temporal information was ambiguous.

In terms of temporal coincidence, the likelihood of perceiving an auditory and visual stimulus as being synchronous is maximized when the auditory stimulus is presented shortly after the visual stimulus, although the specific temporal parameters vary with the specific stimuli being used (Dixon and Spitz, 1980; Lewald and Guski, 2003; Soto-Faraco and Alsisus, 2009; Spence and Squire, 2003; Van Wassenhove *et al.*, 2007). In addition to temporal coincidence, cross-modal stimulus congruency factors play a role in setting the point of subjective synchrony (PSS). Spence (2011) discusses various types of stimulus congruency, including statistical, structural, and semantically mediated relationships. While acknowledging the potential contribution of statistical (e.g., Evans and Treisman, 2009; Marks, 1987) and semantically mediated (e.g., Van Atteveldt *et al.*, 2004; Van der Burg and Goodbourn, 2015; Van der Burg *et al.*, 2010) cross-modal congruency, for the purposes of the current research, the focus will be on structural congruency, represented by the size of visual stimuli and the intensity of auditory stimuli. This cross-modal pairing is based on A Theory of Magnitude (Walsh, 2003), which holds that auditory intensity and visual size serve as common indices of magnitude, such that loud and large stimuli both serve to indicate high magnitude, while quiet and small stimuli both indicate low magnitude. In this context, binding can be promoted through use of stimuli that match in terms of magnitude.

Since temporal and stimulus congruency factors have both been shown to play a role in perception of synchrony, the dominance of temporal coincidence in the Wilbiks and Dyson (2013a, b) paradigm merits further investigation. This is especially true given that the paradigm being employed is richer in visual than auditory stimuli (cf. Wilbiks and Dyson, 2013b), and therefore should rely on stimulus congruency factors as more informative than temporal factors (Alais and Burr, 2004). The root of the respective dominance of visual or auditory information in experiments such as this one can be linked to the amount of information that can be gleaned from each stimulus modality. One possible explanation for this is that the original paradigm involved five levels of temporal variation between stimuli, and only two levels of stimulus congruency. This, coupled with the fact that the temporal stimuli were drawn from a rectangular distribution (i.e., each time of presentation was equiprobable), meant that temporal information provided more useful information than stimulus information to provide for perception of synchrony. When the auditory stimulus was temporally aligned with one of the visual stimuli, it provided a highly useful cue for the association between those two stimuli. On the other hand, when the auditory stimulus was not temporally aligned, it provided less useful information regarding association of stimuli, and as such we observed increases in the effectiveness of stimulus congruency factors.

In considering the informative quality of different features, it is important to note the differences between stimulus expectation and stimulus reliability. Stimulus reliability refers to the degree of certainty of information conveyed by a given sensory channel. For example, a visual stimulus presented with an intensity well above the threshold of perception is considered to be more reliable than one that is presented close to threshold (Bresciani and Ernst, 2007). To this end, perceivers tend to use weighting of information in a binding process that is based on the reliability of the information, such that as one signal gives more clear (i.e., invariant) information, it is assigned a greater weight than a signal that is less reliable (Oruc *et al.*, 2003). Reliability of stimuli has been shown to be increased when a stimulus from a different modality is presented in synchrony with a potential target stimulus (e.g., Van der Burg *et al.*, 2008), and comparative work with macaques has shown that this reliability effect can be shown in single neurons within the brain (Morgan *et al.*, 2008).

Stimulus expectation, however, refers to a combination of the rhythmicity and predictability of stimuli that allow for an individual to successfully detect a stimulus that may otherwise not have been reliable (Ten Oever *et al.*, 2014). Rhythmicity refers to stimuli being presented at standard oscillatory timepoints, which can increase attention and performance in the manner of the Dynamic Attending Theory (DAT; Large and Jones, 1999). However, the current research features a manipulation of the temporal expectation of stimuli being presented in the same manner as Rimmele *et al.* (2011), who found that increases in temporal predictability of stimulus information led to improvement of reaction times as well as concomitant increases in P1 magnitude. In the current research, we manipulate stimulus expectation by varying the distribution from which the timing of the auditory stimulus is drawn. In doing so, we present conditions under which temporal stimulus expectation is either highly predictable, or less predictable.

In a recent series of experiments, Van der Burg *et al.* (2014) manipulated the relative contributions of stimulus information (operationalized by the number of visual stimuli present in a display) and temporal information. They also manipulated the level of ‘clutter’ present in both temporal and visual fields by increasing or decreasing the number of stimuli and the number of timepoints that are used in a given experimental trial. They found that the amount of visual clutter present does not have a significant effect on one’s window of audiovisual simultaneity, but that a greater number of temporal timepoints (40 ms intervals rather than 80 ms intervals) led to a higher degree of difficulty. Van der Burg *et al.* (2014) attribute this difference to perceptual confusion caused by overlap in temporal windows of integration. In considering this result in light of Wilbiks and Dyson (2013a), we can see that having a greater number of timepoints for auditory stimulus presentation that are aligned with a visual stimulus increases the utility of temporal information by increasing participants’ expectation of when it will be presented (as per Rimmele *et al.*, 2011), while presenting a greater number of auditory stimuli that are not aligned with visual stimuli decreases the utility of temporal information.

The current experiment examines whether the relative weighting of temporal presentations can alter the interaction between temporal and stimulus factors. It investigates the effects of drawing from non-rectangular temporal distributions to make temporal information relatively more or less utile in perception. Two non-rectangular temporal distributions are employed; namely, a centrally peaked distribution and a distribution with peaks at extremely long and short SOAs (see Section 2 for full details). We expected to find that temporal information will be more valuable (and thus more likely to be employed) when it provides more definitive information, as will be the case when a peripheral distribution provides more definite timepoints that overlap with anchor stimuli. Alternately, when a centrally-peaked distribution is used, making temporal information less informative, we expect participants will be more likely to use stimulus information to make a decision.

2. Methods

2.1. Participants

Recruitment and research practices were approved by the Research Ethics Board at Mount Allison University. Informed consent from 37 participants was obtained prior to the experiment. Five participants were excluded for responding randomly, not only failing to correlate with time as the majority of the participants did, but also showing no modulation in their results. This is an occasional problem found when recruiting potentially unmotivated undergraduate students who may only attend to earn a class credit, and responding randomly in order to complete the task. The 32 participants making up the final sample had a mean age of 19.9 years (SD = 3.9) and included 22 females and 31 right-handed individuals. All participants self-reported normal or corrected to-normal vision and hearing.

2.2. *Stimuli and Apparatus*

1-kHz sounds of 100 ms duration, with 5 ms linear on-set and off-set ramps were created using SoundEdit 16 (MacroMedia). All sounds were played binaurally from free-field speakers positioned either side of a computer monitor viewed approximately 57 cm away to encourage magnitude coincidence between auditory and visual signals (Calvert *et al.*, 2004). All sounds were calibrated using a Scosche SPL100 sound level meter to approximately 56 or 71 dB(C) to represent quiet and loud sounds, respectively. The visual stimulus consisted of a yellow asterisk, presented in the center of a black screen in either 24- or 96-point Chicago font to represent small and large sizes, as per Wilbiks and Dyson (2013a). Stimulus presentation was controlled by Superlab 5.0, with responses given *via* a computer keyboard.

2.3. *Design and Procedure*

Experimental blocks of 640 trials were developed for each condition involving the orthogonal combination of first visual stimulus magnitude (V1; small, large), auditory stimulus magnitude (A; quiet, loud) and, second visual stimulus magnitude (V2; small, large). These eight (2 × 2 × 2) sets of stimuli were further varied by changing the temporal presentation of the auditory stimulus with respect to the first and second visual stimuli (see Fig. 1). The two visual stimuli were always presented with a 200 ms interstimulus interval (ISI), and were presented for 100 ms each. The auditory stimulus could occur simultaneously with the onset of the first visual stimulus (0 ms), and at 100 ms intervals until 100 ms after the onset of the second visual stimulus (400 ms), for a total of five possible temporal presentations. These timepoints were chosen for the onset of the to-be-bound stimulus to coincide with the respective onsets (0 and 300 ms) and offsets (100 and 400 ms) of each anchor, as well as fifth timepoint (200 ms) where only the to-be-bound offset and second anchor on-set were associated. The critical manipulation of distribution was operationalized by presenting stimuli drawn from one of three temporal distributions (see Fig. 1): rectangular (0, 100, 200, 300, 400 ms SOA each occurring 16 times per stimulus condition), peripheral (0 ms = 25 trials,

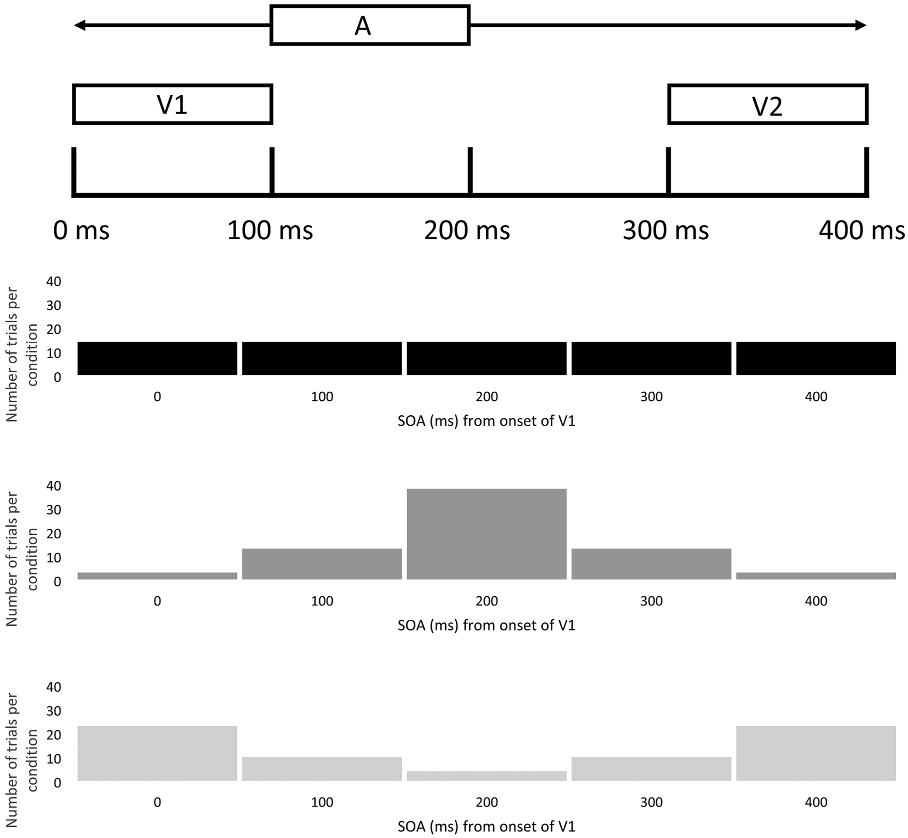

Figure 1. Schematic of experimental procedure. Upper panel shows the potential timings for visual and auditory stimuli. Lower panel shows three temporal distributions from which auditory stimulus timings were drawn.

100 ms = 12, 200 ms = 6, 300 ms = 12, 400 ms = 25), and central (0 ms = 5 trials, 100 ms = 15, 200 ms = 40, 300 ms = 15, 400 ms = 5). Participants each completed a rectangular block and one of the other blocks (peripheral or central), in a counter-balanced order.

Each trial began with the presentation of a blank screen for 500 ms, followed by a variable lag. Participants were then presented with the first visual stimulus for 100 ms, and following a 200 ms interval the second visual stimulus was presented for 100 ms. The to-be-bound auditory stimulus was presented at some time between the two visual stimulus presentations, at 100 ms intervals. Stimulus presentation was followed, 500 ms after the offset of the second visual stimulus, by a response prompt saying ‘FIRST OR SECOND?’ Participants were asked to respond by pressing F on a standard keyboard if they thought the to-be-bound stimulus was caused by the first anchor, and J if they thought the to-be-bound stimulus was caused by the second anchor. As a result of the subjective nature of the task, no feedback was provided.

3. Results and Discussion

The data were separated into two groups for the purposes of analysis, such that there were 16 participants who had completed a central block with a rectangular block, and another 16 participants who had completed a peripheral block with a rectangular block. Proportions for second stimulus responding were subjected to a five-way repeated measures ANOVA with the factors of distribution (central/peripheral, rectangular) × first visual stimulus [V1] (small, large) × auditory stimulus [A] (quiet, loud) × second visual stimulus [V2] (small, large) × time (0, 100, 200, 300, 400 ms). These analyses were conducted separately for comparing rectangular to central distributions (data are displayed in Fig. 2; full results of ANOVA are displayed in Table 1), and for comparing rectangular to peripheral (data are displayed in Fig. 3; full results of ANOVA are displayed in Table 2) with informative effects summarized below.

In comparing rectangular distributions to central distributions a significant effect of first visual stimulus magnitude ($p = 0.002$) revealed a greater tendency for secondary responding when the first visual stimulus was small,

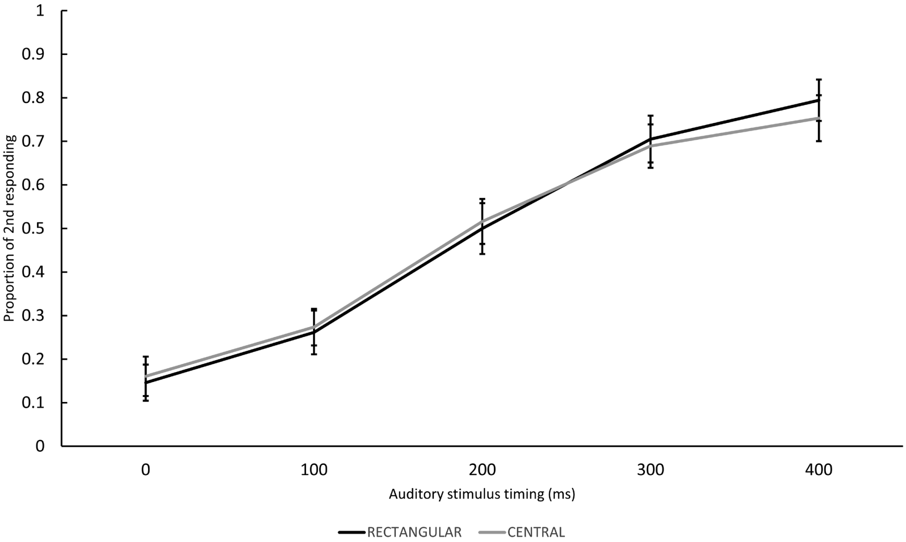

Figure 2. Proportion of perceived synchrony between auditory stimulus and second visual stimulus based on timing of presentation and comparing rectangular and centrally peaked distributions of temporal presentation. Error bars represent standard error.

Table 1.
Results of 2 × 2 × 2 × 2 × 5 ANOVA on proportion second responding for comparison of rectangular distributions to central distributions. Bold text indicates significant results

Measure	df	F	MSE	p	η_p^2
Distribution (D)	1, 15	0.024	0.106	0.878	0.002
First Visual (V1)	**1, 15**	**14.109**	**0.150**	**0.002**	**0.485**
Auditory (A)	**1, 15**	**6.543**	**0.013**	**0.022**	**0.304**
Second Visual (V2)	1, 15	1.913	0.055	0.187	0.113
Time (T)	**4, 60**	**72.786**	**0.252**	**<0.001**	**0.829**
D × V1	1, 15	0.170	0.032	0.686	0.011
D × A	1, 15	0.219	0.007	0.646	0.014
D × V2	1, 15	0.667	0.010	0.427	0.043
D × T	4, 60	1.974	0.021	0.110	0.116
V1 × A	1, 15	3.713	0.025	0.073	0.198
V1 × V2	1, 15	2.347	0.011	0.146	0.135
V1 × T	**4, 60**	**6.282**	**0.027**	**<0.001**	**0.295**
A × V2	1, 15	3.379	0.015	0.086	0.184
A × T	4, 60	0.393	0.017	0.813	0.026
V2 × T	4, 60	0.296	0.012	0.879	0.019
D × V1 × A	1, 15	1.085	0.005	0.314	0.067
D × V1 × V2	1, 15	0.031	0.008	0.863	0.002
D × V1 × T	**4, 60**	**3.060**	**0.008**	**0.023**	**0.169**
D × A × V2	1, 15	3.179	0.013	0.095	0.175
D × A × T	4, 60	0.848	0.013	0.500	0.054
D × V2 × T	4, 60	0.154	0.010	0.960	0.010
V1 × A × V2	1, 15	0.871	0.020	0.365	0.055
V1 × A × T	4, 60	2.345	0.019	0.065	0.135
V1 × V2 × T	4, 60	2.364	0.020	0.063	0.136
A × V2 × T	4, 60	1.448	0.015	0.229	0.088
D × V1 × A × V2	1, 15	0.030	0.008	0.864	0.002
D × V1 × A × T	4, 60	0.405	0.009	0.804	0.026
D × V1 × V2 × T	4, 60	1.570	0.008	0.194	0.095
D × A × V2 × T	4, 60	0.199	0.005	0.938	0.013
V1 × A × V2 × T	4, 60	0.123	0.015	0.974	0.008
D × V1 × A × V2 × T	4, 60	0.353	0.008	0.841	0.023

while a significant effect of auditory stimulus magnitude ($p = 0.022$) revealed a greater tendency for primary responding when the auditory stimulus was quiet. An effect of time ($p < 0.001$), as expected, revealed an increase in secondary responding with later presentation. This main effect was subsumed by a first visual stimulus × time interaction ($p < 0.001$), and Tukey's HSD (honest significant difference) ($p < 0.05$) tests revealed that the effect of first visual stimulus magnitude was only significant at SOAs of 200, 300, and 400 ms, when the presentation of the auditory stimulus did not overlap

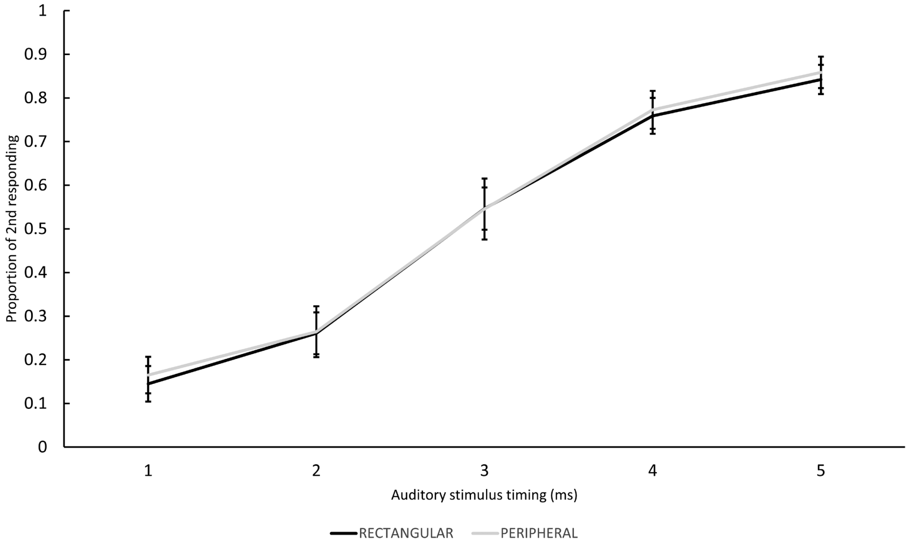

Figure 3. Proportion of perceived synchrony between auditory stimulus and second visual stimulus based on timing of presentation and comparing rectangular and peripherally peaked distributions of temporal presentation. Error bars represent standard error.

with the first visual stimulus. In comparing rectangular and peripheral distributions, similar findings showed an effect of first visual stimulus ($p < 0.001$), time ($p < 0.001$), and a significant first visual stimulus × time interaction ($p = 0.024$). There were some additional significant higher order interactions (all displayed in Tables 1 and 2: Central: D × V1 × T ($p = 0.023$), Peripheral: V1 × A × T ($p = 0.047$), V1 × A × V2 × T ($p = 0.040$)), but Tukey's HSD ($p < 0.05$) did not reveal any significant differences within the interactions.

While these results replicate the findings from Wilbiks and Dyson (2013a, b), the question of interest in the present research is how the relative weighting of temporal and stimulus factors is determined in the competitive decision-making process. To that end, a comparison was made between proportion of secondary responding in rectangular *vs.* central and in rectangular *vs.* peripheral distributions. It was expected that when temporal information was of increased informational value (i.e., peripheral distribution) participants would be more likely to assign temporal information greater weight than stimulus congruency information. Alternately, when temporal information was of decreased informational value (i.e., central distribution) participants would assign it a lesser weight than stimulus congruency information. This was operationalized in two ways within this experiment, both of which were predicated on a finding demonstrated by Roseboom *et al.* (2009). They show that once a cross-modal binding is made, that bind is very difficult to break, regardless of additional information that is presented *post hoc*.

Table 2.
Results of 2 × 2 × 2 × 2 × 5 ANOVA on proportion second responding for comparison of rectangular distributions to peripheral distributions. Bold text indicates significant results

Measure	df	F	MSE	p	η_p^2
Distribution (D)	1, 15	0.270	0.134	0.611	0.018
First Visual (V1)	**1, 15**	**25.645**	**0.072**	**<0.001**	**0.631**
Auditory (A)	1, 15	1.515	0.019	0.237	0.092
Second Visual (V2)	1, 15	3.895	0.042	0.067	0.206
Time (T)	**4, 60**	**98.318**	**0.241**	**<0.001**	**0.868**
D × V1	**1, 15**	**6.963**	**0.005**	**0.019**	**0.317**
D × A	1, 15	3.296	0.010	0.089	0.180
D × V2	1, 15	2.756	0.010	0.118	0.155
D × T	4, 60	0.090	0.055	0.985	0.006
V1 × A	**1, 15**	**6.162**	**0.069**	**0.025**	**0.291**
V1 × V2	1, 15	1.251	0.021	0.281	0.077
V1 × T	**4, 60**	**3.043**	**0.023**	**0.024**	**0.169**
A × V2	**1, 15**	**11.626**	**0.022**	**0.004**	**0.437**
A × T	4, 60	2.482	0.011	0.053	0.142
V2 × T	4, 60	0.100	0.013	0.982	0.007
D × V1 × A	1, 15	0.001	0.010	0.990	0.001
D × V1 × V2	1, 15	0.275	0.015	0.608	0.018
D × V1 × T	4, 60	0.965	0.012	0.434	0.060
D × A × V2	1, 15	0.065	0.024	0.802	0.004
D × A × T	4, 60	0.716	0.012	0.585	0.046
D × V2 × T	4, 60	2.080	0.015	0.095	0.122
V1 × A × V2	1, 15	0.190	0.019	0.669	0.012
V1 × A × T	**4, 60**	**2.574**	**0.011**	**0.047**	**0.146**
V1 × V2 × T	4, 60	0.280	0.018	0.890	0.018
A × V2 × T	4, 60	1.923	0.010	0.118	0.114
D × V1 × A × V2	1, 15	0.055	0.014	0.818	0.004
D × V1 × A × T	4, 60	0.234	0.012	0.918	0.015
D × V1 × V2 × T	4, 60	0.428	0.014	0.788	0.028
D × A × V2 × T	4, 60	0.583	0.013	0.676	0.037
V1 × A × V2 × T	**4, 60**	**2.673**	**0.008**	**0.040**	**0.151**
D × V1 × A × V2 × T	4, 60	0.200	0.015	0.938	0.013

In a paradigm such as the current one, visual and auditory information are presented sequentially, which means that the first visual stimulus has an advantage over the second visual stimulus in that it is presented first (as per Roseboom *et al.*, 2009). Once that bind with the first visual stimulus is made, it is difficult to break the bind, even if later information regarding stimulus congruency should lead to a different decision. To this end, I obtained an index of release from the first visual anchor by extracting two values from secondary responding rates after fitting those data to a psychophysical function: the point

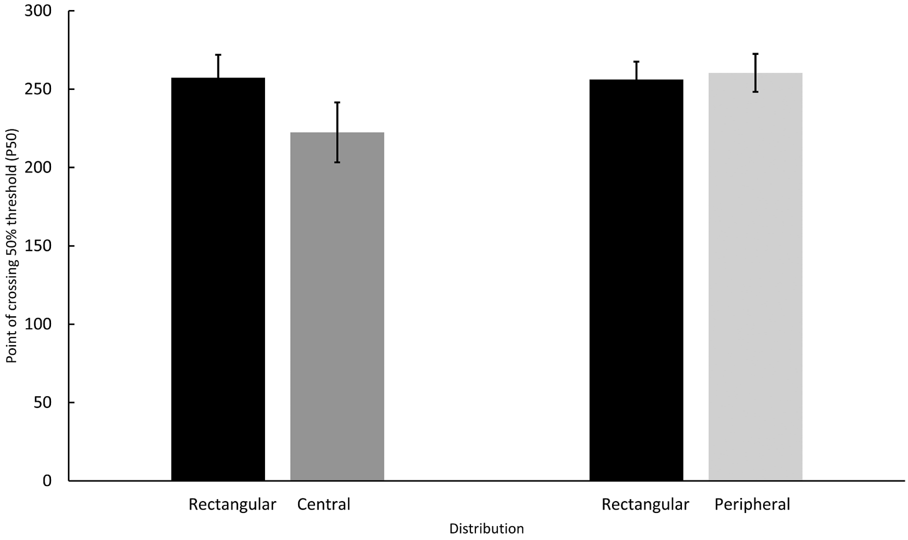

Figure 4. Point of crossing 50% secondary responding (P50) for comparing rectangular and centrally peaked distributions (left side of figure) and for comparing rectangular and peripherally peaked distributions (right side of figure). Error bars represent standard error.

at which responding crossed a 50% threshold (P50), and the tangential slope (m_T) at that point. The P50 (see Fig. 4) serves as a measure of how long participants tended towards primary responding before being released from the first visual stimulus to enable potential binding to the second visual stimulus. Similarly, but not identically, the m_T (see Fig. 5) is indicative of both the difference between minimal and maximal responding (an index of the strength of weight assigned to temporal factors), as well as the length of time a participant remained likely to respond to the first visual stimulus. Lower scores for P50 and m_T suggest higher relative weighting given to stimulus congruency factors, as stimulus congruency enabled release from the first visual stimulus at an earlier time, while higher scores for both P50 and m_T suggest higher relative weighting given to temporal factors, and a lack of release from the first visual stimulus.

These measures were submitted to paired-sample t-tests for the comparison of rectangular to central, and of rectangular to peripheral distributions. For comparing rectangular to central, P50 was found to be significantly higher for rectangular distributions than central distributions, $t(15) = 2.224$, $p = 0.042$, Cohen's $d = 0.512$. Similarly, m_T was greater for rectangular than central, $t(15) = 2.338$, $p = 0.034$, Cohen's $d = 0.437$. These findings indicate that one is more likely to use temporal factors when drawing from a rectangular distribution than when drawing from a central distribution, as was predicted. With a central distribution, a dearth of informative temporal information

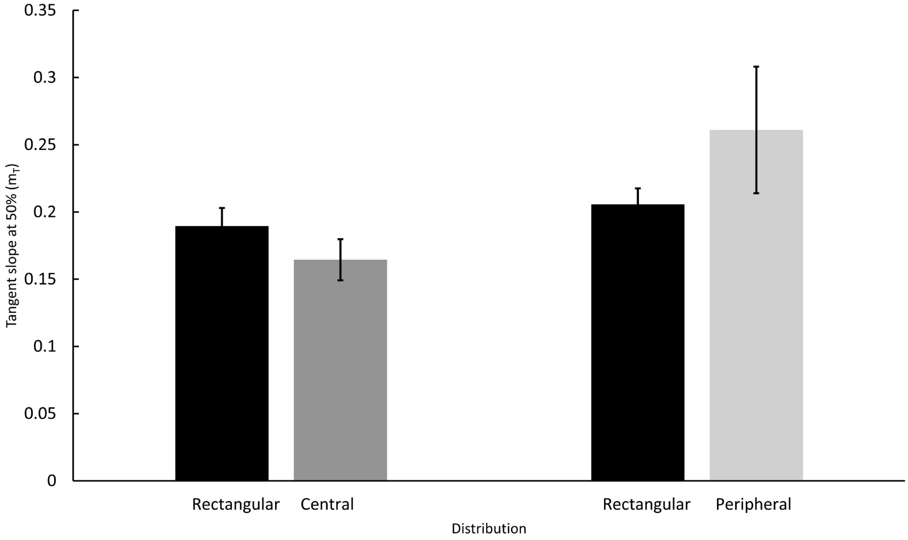

Figure 5. Tangent slope at P50 (m_T) for comparing rectangular and centrally peaked distributions (left side of figure) and for comparing rectangular and peripherally peaked distributions (right side of figure). Error bars represent standard error.

makes it a less useful factor to rely on, and as such individuals tend to use stimulus information instead. When comparing rectangular to peripheral distributions, no such differences were found [for P50, $t(15) = -0.513$, $p = 0.615$, Cohen's $d = 0.090$; for m_T, $t(15) = -1.101$, $p = 0.288$, Cohen's $d = 0.198$]. So the data suggest that increasing the informative value of temporal information by using a peripheral distribution does not have a similar effect to that found in decreasing informative value (although the numerical difference is in the expected direction).

When dealing with studies of crossmodal interaction, it is important to consider the potential for confound between true cross-modal effects and response biases. One could analyse response bias — the degree to which participants' responses are dictated by their implicit decision-making criteria — by using signal detection theory methodologies and comparing the ratio of correct detections to false alarms (e.g., Marks *et al.*, 2003). However, this type of analysis relies on the existence of a 'correct' response, and in the present research there is no correct or incorrect response. However, it is unlikely that response bias is responsible for the critical effect of temporal distribution in the current research, as response bias is more likely to affect individual stimulus characteristics (e.g., stimulus intensity — Marks *et al.*, 1986) rather than its temporal distribution.

This experiment was conducted to ascertain whether altering the informational value of temporal information through expectation (Rimmele *et al.*,

2011) affects responding in a competitive audiovisual perception task (Wilbiks and Dyson, 2013a, b). It was found that, relative to a baseline rectangular temporal distribution, decreasing the informative value of temporal information leads to a decrease in reliance on such information and a corresponding increase in reliance on stimulus information as occurs naturally in a spatial task (Alais and Burr, 2004; Welch and Warren, 1980). Alternately, increasing the informative value of temporal information does not lead to a significant increase in reliance on temporal information. This asymmetry was not expected, but is also not unprecedented within the study of audiovisual integration (cf. Van Wassenhove *et al.*, 2007; Wada *et al.*, 2003; Wilbiks and Dyson, 2013b). For example, audiovisual perceived synchrony is subject to a greater influence of recalibration when there are more auditory than visual stimuli, as compared to a lack of malleability when there are more visual than auditory stimuli (Wilbiks and Dyson, 2013b). The fact that responding is less reliant on temporal information when temporal information is less informative is reminiscent of the modality appropriateness hypothesis (Welch and Warren, 1980), in that the perceptual system is able to provide preferential weighting to stimuli that provide more useful information. While it is clear in this case that temporal information is carrying less weight in the decision-making process, what is not indicated by the current data is whether this re-weighting is necessarily coupled with an increase in the weight allotted to stimulus information. In future research, it would be of interest to operationalize the use of stimulus information such that they can be examined. For the time being, however, the evidence shows that when temporal information is reduced in utility, perceivers tend to reduce the value they attach to it.

References

Alais, D. and Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration, *Curr. Biol.* **14**, 257–262.

Bresciani, J.-P. and Ernst, M. O. (2007). Signal reliability modulates auditory-tactile integration for event counting, *Neuroreport* **18**, 1157–1161.

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.

Calvert, G. A., Spence, C. and Stein, B. E. (2004). *The Handbook of Multisensory Processing*. MIT Press, Cambridge, MA.

Dixon, N. F. and Spitz, L. (1980). The detection of auditory visual desynchrony, *Perception* **9**, 719–721.

Ernst, M. O. and Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion, *Nature* **415**(6870), 429–433.

Ernst, M. O. and Bulthoff, H. H. (2004). Merging the senses into a robust percept, *Trends Cogn. Sci.* **8**, 162–169.

Evans, K. K. and Treisman, A. (2009). Natural cross-modal mappings between visual and auditory features, *J. Vis.* **10**, 6. DOI:10.1167/10.1.6.

Large, E. W. and Jones, M. R. (1999). The dynamics of attending: how people track time-varying events, *Psychol. Rev.* **106**, 119–159.

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.

Marks, L. E. (1987). On cross-modal similarity: auditory-visual interactions in speeded discrimination, *J. Exp. Psychol. Hum. Percept. Perform.* **13**, 384–394.

Marks, L. E., Szczesiul, R. and Ohlott, P. (1986). On the cross-modal perception of intensity, *J. Exp. Psychol. Hum. Percept. Perform.* **12**, 517–534.

Marks, L. E., Ben-Artzi, E. and Lakatos, S. (2003). Cross-modal interactions in auditory and visual discrimination, *Int. J. Psychophysiol.* **50**, 125–145.

Morgan, M. L., DeAngelis, G. C. and Angelaki, D. E. (2008). Multisensory integration in macaque visual cortex depends on cue reliability, *Neuron* **59**, 662–673.

Oruc, I., Maloney, L. T. and Landy, M. S. (2003). Weighted linear cue combination with possibly correlated error, *Vis. Res.* **43**, 2451–2468.

Rimmele, J., Jolsvai, H. and Sussman, E. (2011). Auditory target detection is affected by implicit temporal and spatial expectations, *J. Cogn. Neurosci.* **23**, 1136–1147.

Roseboom, W., Nishida, S. and Arnold, D. H. (2009). The sliding window of audio-visual simultaneity, *J. Vis.* **9**, 4. DOI:10.1167/9.12.4.

Soto-Faraco, S. and Alsius, A. (2009). Deconstructing the McGurk–MacDonald illusion, *J. Exp. Psychol. Hum. Percept. Perform.* **35**, 580–587.

Spence, C. (2011). Crossmodal correspondences: a tutorial review, *Atten. Percept. Psychophys.* **73**, 971–995.

Spence, C. and Squire, S. (2003). Multisensory integration: maintaining the perception of synchrony, *Curr. Biol.* **13**, R519–R521.

Ten Oever, S., Schroeder, C. E., Poeppel, D., Van Atteveldt, N. and Zion-Golumbic, E. (2014). Rhythmicity and cross-modal temporal cues facilitate detection, *Neuropsychologia* **63**, 43–50.

Van Atteveldt, N., Formisano, E., Goebel, R. and Blomert, L. (2004). Integration of letters and speech sounds in the human brain, *Neuron* **43**, 271–282.

Van der Burg, E. and Goodbourn, P. T. (2015). Rapid, generalized adaptation to asynchronous audiovisual speech, *Proc. R. Soc. Lond. B Biol. Sci.* **282**(1804), 20143083. DOI:10.1098/rspb.2014.3083.

Van der Burg, E., Olivers, C. N. L., 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.

Van der Burg, E., Brederoo, S. G., Nieuwenstein, M. R., Theeuwes, J. and Olivers, C. N. L. (2010). Audiovisual semantic interference and attention: evidence from the attentional blink paradigm, *Acta Psychol.* **134**, 198–205.

Van der Burg, E., Cass, J. and Alais, D. (2014). Window of audio-visual simultaneity is unaffected by spatio-temporal visual clutter, *Sci. Rep.* **4**, 5098. DOI:10.1038/srep05098.

Van Wassenhove, V., Grant, K. W. and Poeppel, D. (2007). Temporal window of integration in auditory-visual speech perception, *Neurophysiologica* **45**, 598–607.

Wada, Y., Kitagawa, N. and Noguchi, K. (2003). Audio-visual integration in temporal perception, *Int. J. Psychophysiol.* **50**, 117–124.

Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space, and quantity, *Trends Cogn. Sci.* **7**, 483–488.

Welch, R. B. and Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy, *Psychol. Bull.* **88**, 638–667.

Wilbiks, J. M. P. and Dyson, B. J. (2013a). Effects of temporal asynchrony and magnitude congruency on competitive audio-visual binding, *Atten. Percept. Psychophys.* **75**, 1883–1891.

Wilbiks, J. M. P. and Dyson, B. J. (2013b). The influence of previous environmental history on audio-visual binding occurs during visual-weighted but not auditory-weighted environments, *Multisens. Res.* **26**, 561–568.