#

**Perceived Audio-Visual Simultaneity is Recalibrated by the Visual Intensity of the Preceding Trial**

**Ryan Horsfall1, Neil Harrison2, Georg Meyer1 and Sophie Wuerger1**

1 Institute of Population Health, University of Liverpool, Waterhouse Building, Block B, Brownlow Street, Liverpool, L69 3GF, United Kingdom

2 Department of Psychology, Liverpool Hope University, Liverpool, L16 9JD, United Kingdom

\* Corresponding author; e-mail: Ryan.horsfall.uk@gmail.com

**Abstract**

A vital heuristic used when making judgements on whether audio-visual signals arise from the same event, is the temporal coincidence of the respective signals. Previous research has highlighted a process, whereby the perception of simultaneity rapidly recalibrates to account for differences in the physical temporal offsets of stimuli. The current paper investigated whether rapid recalibration also occurs in response to differences in central arrival latencies, driven by visual-intensity-dependent processing times. In a behavioural experiment, observers completed a temporal-order judgement (TOJ), simultaneity judgement (SJ) and simple reaction-time (RT) task and responded to audio-visual trials that were preceded by other audio-visual trials with either a bright or dim visual stimulus. It was found that the point of subjective simultaneity shifted, due to the visual intensity of the preceding stimulus, in the TOJ, but not SJ task, while the RT data revealed no effect of preceding intensity. Our data therefore provide some evidence that the perception of simultaneity rapidly recalibrates based on stimulus intensity.

**Keywords**

Audio-visual, rapid recalibration, perceptual judgement

1. **Introduction**

Neural processing latencies of sensory signals depend on many factors (e.g. Gondan *et al.*, 2005; Leone and McCourt, 2012, 2013), including stimulus modality and intensity (Jaśkowski *et al.*, 1995; Pöppel *et al.*, 1990; Ulrich *et al.*, 1998). Auditory signals typically reach the respective sensory cortex at shorter latencies than visual stimuli (Raij et al., 2010); however, an intensity dependency is evident for both the visual and auditory modalities, with shorter processing latencies for higher-intensity stimuli (Carrillo-de-la-Pena et al., 1999; Jaskownki et al., 1994; Plainis et al., 2013). One of the consequences of this intensity dependence is that the perceived simultaneity of an audio-visual event depends on the relative intensities of the unimodal signals (Horsfall *et al.*, 2021a; Leone and McCourt, 2015; Roufs, 1963).

* 1. *Recalibration of Perceived Simultaneity*

It has been argued that humans recalibrate their audio-visual integration mechanisms to deal with the different arrival times of auditory and visual signals (Fujisaki *et al.*, 2004), thereby achieving a perception of simultaneity which is, to a large extent, independent of stimulus-related processing properties. It has been shown that this recalibration occurs in response to prolonged exposure to temporally offset stimuli (Fujisaki et al., 2004), as well as on a dynamic, trial-to-trial basis (Noel et al., 2016; Simon et al., 2017; Van der Burg et al., 2013, 2014).

* 1. *Proposed Intensity-Dependent Compensatory Mechanism*

An ideal audio-visual integrator should be independent of intensity for signals emanating from a fixed distance from the observer, as intensity variations in processing latencies do not reflect the physical timing properties of stimuli in real-world scenarios. Since intensity differences are mapped into neural time delays early in the processing pathway, we hypothesise that rapid recalibration could play a role in minimising the intensity dependence of perceived audio-visual simultaneity. In other words, we hypothesise that rapid recalibration is modulated by the relative, intensity-dependent central arrival times of audio-visual signals, rather than solely the temporaloffset of these signals. Since low-intensity visual signals reach cortical processing areas at longer latencies, we predict that this processing delay will be compensated for as if it were a physical stimulus offset, resulting in a shift in the point of subjective simultaneity (PSS) (see Fig. 1), which reflects an estimate of the temporal offset at which the two signal are perceived as simultaneous.

* 1. *Task-Specific Processes*

The simultaneity judgement (SJ) and temporal-order judgement (TOJ) tasks are commonly used to assess the perceived relative timing of audio-visual events (e.g. Yarrow *et al.*, 2011; Zampini *et al.*, 2003, 2005). The SJ and TOJ tasks both provide estimates of the PSS, though previous research often fails to find a significant correlation between these estimates from the two tasks (Basharat *et al.*, 2018; Binder, 2015; Linares and Holcombe, 2014; Love *et al.*, 2013; Van Eijk *et al.*, 2008; Vatakis *et al.*, 2008). Furthermore, the effect of rapid recalibration is reportedly inconsistent between the SJ and TOJ tasks (Roseboom, 2019). Various models have proposed shared sensory processing when making judgements on simultaneity and temporal order, with differences at the decision stage (García-Pérez and Alcalá-Quintana, 2012; Jaśkowski, 1991; Parise and Ernst, 2016; Stelmach and Herdman, 1991).

* 1. *The Current Study*

The main aims of the study were to: (a) investigate whether rapid recalibration is modulated by the relative, intensity-dependent central arrival times of audio-visual signals, and (b) investigate whether this proposed shift in perceived simultaneity is reflected by changes in relative processing latencies, as evident in simple RTs (e.g. Pins and Bonnet, 1997).

1. **Method**
2. *Participants*

The behavioural study involved 21 participants, aged 19–49 (mean = 27.10, SD = 7.80, median = 25), nine of whom were female and 13 of whom were naïve to the SJ and TOJ tasks. All observers reported normal or corrected-to-normal vision and hearing, and no neurological disorders. Opportunity sampling was used to recruit participants through advertisements delivered to staff and students across the University of Liverpool via announcement e-mails. The PSSs of four participants were estimated to be outside of the testable range in at least one of the conditions of this analysis, and they were therefore removed from all analyses.

1. *Design*

A repeated-measures design was used. Participants completed the audio-visual SJ task, TOJ task and a simple reaction-time (RT) task on three separate days. The order of the TOJ and SJ tasks was counterbalanced, and the RT task was always completed second.

1. *Apparatus*

Participants were seated in a soundproof booth (IAC, Winchester, UK), 113 cm from a LED and speaker that were held at roughly eye level in an adjustable clamp. A Tucker Davies RP2.1 real-time processor (TDT technologies, Alachua, FL, USA) was used to generate the visual and auditory stimuli. A single speaker of a Xenta M-219 Notebook (Xenta, East Riding of Yorkshire, United Kingdom) produced the auditory stimuli. This was located 1.62° below a 5-mm white LED. A custom-built button box, consisting of two wide, shallow buttons was used to record the participant’s responses. MATLAB was used on a PC located outside of the soundproof booth to signal the Tucker Davies system and record responses.

1. *Preliminary Threshold Estimation*

The visual and auditory stimuli were chosen by estimating the visual and auditory thresholds of 80 participants (age range: 22–43, mean = 27.00, SD = 2.39, median = 24.50) in a pilot experiment (for full procedure, see Horsfall et al., 2021b). After 15 minutes of adaptation to the dark, participants completed two separate two-alternative forced choice (2AFC) tasks, consisting of a visual threshold estimation task, followed by an auditory threshold estimation task. The threshold estimation used a QUEST (Watson and Pelli, 1983) procedure as this is more efficient than the typical threshold estimation involving sequential testing (Watson and Pelli, 1983). The mean visual threshold was 0.005 cd/m2 and the mean auditory threshold was calculated as 20.54 dB.

1. *Stimuli*

The stimuli consisted of a 100-ms visual flash and a 100-ms auditory beep, that were presented at consistent intensities for each participant. The visual flash was presented at three intensities (0.02 cd/m2, 0.16 cd/m2 and 1.34 cd/m2) which were 6 dB, 15 dB and 24 dB above the previously estimated mean visual threshold (see Section 2.4). The auditory beep had a fixed frequency of 1000 Hz, and was presented with a flat amplitude envelope. The intensity of the auditory stimulus was 15 dB above the previously estimated threshold (see Section 2.4) in an attempt to make sure the intensity was comparable to that of the visual stimuli. The inter-stimulus interval was 1500 ms, plus a random value between 0 and 2000 ms.

1. *Behavioural Trials*

Stimuli were identical across the SJ, TOJ and RT tasks. The audio-visual stimuli were presented at seven stimulus onset asynchronies (SOAs) (−200 ms, −100 ms, −50 ms, 0 ms, 50 ms, 100 ms, 200 ms). The three intensities were presented so that every second trial (aside from the filler trials) was of the middle (0.16 cd/m2) intensity (see Fig. 1), which we define as the ‘test’ intensity. Within each block, a trial at each of the seven SOAs of the test intensity was preceded by a trial at each of the seven SOAs of the dim (0.02 cd/m2)and bright (1.34 cd/m2) intensities, such that each of these SOAs were presented in random order. Therefore, there were 98 test trials in each block, which were preceded by 49 dim trials and 49 bright trials, plus 32 ‘filler’ trials. These consisted of two audio-visual trials with random flash intensities (either bright or dim), each presented with a random offset between (−200 and 200 ms), followed by a test intensity trial. These were included to ensure participants did not become aware of the pattern of stimulus presentations. All filler trials, alongside the following test intensity trials were not included in any analyses. Both the SJ and TOJ tasks consisted of seven blocks each. Over the course of these seven blocks were a total of 98 trials at each SOA of the test intensity and 49 trials at each SOA in the dim and bright conditions, totalling 1596 trials in each task (including pairs of filler trials). In the RT task, participants completed four blocks of trials which are identical to those presented in the SJ and TOJ tasks, summing to 488 trials in total.

1. *Data Fitting*

To estimate the PSS in the TOJ task, a sigmoidal curve (Boltzmann) was fitted to the relative frequency of ‘light first’ responses:

 (*a* − *b*)/{1 + *exp*[(*x* − *c*)/*d*]} + *b* (1)

where *a* reflects the upper asymptote, and *b* reflects the lower asymptote. The parameter *d* represents the width of the curve, whereas *c* was the location of the curve. The SOA (in ms) corresponding to a 50% response rate of ‘light-first’ button presses was used as an estimate of the PSS.

For the SJ data, a Gaussian was fitted to the relative frequency of ‘simultaneous’ responses:

a + b · exp{−0.5 · [(x − c)/d]2} (2)

where *a* reflects the offset along the y-axis, *b,* the amplitude of the curve and *c* reflects the midpoint which was used as an estimate of the PSS. The parameter *d* reflects the width of the curve. A nonlinear least-squares method (function ‘fit’; MATLAB2018) was used to find the best-fitting parameters for both the TOJ and SJ data.

1. *Data Analysis*

A one-way repeated-measures ANOVA was used to assess the effect of visual intensity on the PSS in the SJ and TOJ tasks. Here, a Greenhouse–Geisser correction was applied where the assumption of sphericity was violated. The Pearson correlation coefficient was used to assess the correlation between PSSs recorded in the SJ and TOJ tasks. This was done separately at each of the three intensities.

The test intensity data of the SJ and TOJ tasks were split based on the SOA and visual intensity of the preceding trial, and the resultant relative frequencies (either of ‘light first’ or ‘simultaneous’ responses) were fitted using functions described in Section 2.7. Regression lines were fitted independently to each observer’s PSSs, as a function of preceding SOA. This was done separately for trials preceded by dim or bright trials (i.e. split data), and for trials preceded by both bright and dim trials (i.e. all test intensity trials). It should be noted that while regression lines were used as a best fit for the current data set (see the Appendix), we would expect the relationship between preceding SOA and PSSs to be linear only across a small range of SOAs, and to return to zero at higher (and more noticeable) temporal offsets (see Van der Burg *et al.*, 2013).

Paired *t* tests were then used to compare the intercepts of the regression lines fitted to data preceded by dim or bright trials. One-tailed tests were used, as the prediction for the effect of preceding intensity was directional (see Fig. 2).

For the RT data, a 7 × 3 repeated-measures ANOVA was used to assess the effect of SOA (−200 to 200) and intensity (dim, test intensity, and bright) on reaction times (Simon *et al.*, 2018). A 7 × 2 × 2 repeated-measures ANOVA was then used to assess the effect of SOA, preceding intensity, and preceding leading modality (audio-leading or visual-leading), on RTs at the test intensity. Greenhouse–Geisser corrections were applied where appropriate.

1. *Quantitative Prediction*

Figure 2D displays the quantitative prediction. The predicted effect of preceding intensity (cyan and grey dashed lines), as a function of preceding SOA, was calculated by mapping the intensity-dependent detection delay of the visual stimulus (taken from Horsfall *et al.*, 2021a), into a physical delay. This was done by calculating the differences between the highest/lowest intensity PSSs (taken from Horsfall *et al.*, 2021a) and the predicted mean PSS for the test intensity. These were used to infer the difference in visual processing latencies (i.e., the intensity-dependent detection delay) between the bright/dim stimuli and the test intensity stimuli. This was then factored in when calculating the preceding-SOA-dependent shift; the magnitude of this shift was predicted using the data of Simon et al., (who investigated the effect of preceding SOA on PSS), by taking the size of the preceding-SOA-dependent shifts of the corresponding SOAs of Simon et al.’s (2017) data, and plotting this relative to our predicted test intensity mean PSS.

Since visual processing latencies are intensity-dependent (Lines *et al.*, 1984), visual signals in bright trials (Fig. 2B) will reach multisensory areas of the brain at shorter latencies than dim signals (Fig. 2C). We therefore posit that the offset in these intensity-dependent central arrival times will then be recalibrated in the subsequent trial (Fig. 2D), as if they were physical temporal offsets.

1. *Procedure*

For the SJ task, participants were asked to hold the button box vertically, and to press the top button if they believed that the two stimuli occurred simultaneously and the bottom button if they occurred non-simultaneously. For the TOJ task participants were asked to hold the button box horizontally, and to press the left button if they perceived the flash to have occurred before the beep, and the right button if they perceived the beep first. Participants were requested to hold the button box with different orientations to create a clear distinction between the two tasks, in order to minimise the potential of the participant habitually performing one task, when they should have changed to the other. Participants were asked to respond to every trial and to use their best guess when unsure of the correct response. For the RT task participants were instructed to respond as quickly as possible with a left button press whenever they perceived a light or sound (or both). Participants also completed a unimodal simple RT task, the data from which formed part of a separate experiment. Before both the initial SJ or TOJ block, participants completed two practice blocks of the upcoming task, consisting of 36 trials of the highest visual intensity, split across 3 SOAs (−200 ms, 0 ms and 200 ms) in the SJ and RT task and 2 SOAs (−200 ms and 200 ms) in the TOJ task, in order to increase familiarity with the buttons. Participants would only progress to the main SJ and TOJ task if they achieved at least a 66.66% correct response rate in the respective practice session in one of their two attempts. An opportunity to ask questions was provided after both blocks. Participants were then given 15 minutes of dark adaptation before the experiment began, during which they were asked to relax. Short breaks were provided between each block. After the final (third) session was completed, the participants provided written answers to questions aiming to obtain basic participant information, along with a question on whether the participant had noticed any pattern in the trials/stimuli presented during the experiment. This final question was given with the intention of removing any participant who became aware that the test intensity was always preceded by a dim or bright trial, though no participants reported this.

1. **Results**
2. *Effect of Intensity on Perceived Simultaneity*

One of the main aims of our analysis was to investigate the effect of preceding intensity on PSS. However, we must first ascertain whether our intensity manipulation had a reliable effect on PSS across the two tasks (see Fig. 3). In the TOJ task (Fig. 4, top row), there was a significant main effect of intensity on PSS (*F*1.31,20.99 = 34.51, *p* < 0.001, *η*p*2* = 0.68). Bonferroni-corrected post-hoc tests revealed significant differences between each of the intensities (*p* < 0.01), with higher PSSs (indicating the light had to come earlier, relative to the sound, in order for simultaneity to be perceived) for dimmer stimuli. Likewise, in the SJ task (Fig. 4, bottom row), there was a significant main effect of intensity (*F*1.48,23.71 = 49.00, *p* < 0.001, *η*p*2* = 0.75), and post-hoc tests revealed significant differences between the PSSs at each intensity (*p* < 0.01), with higher PSSs for dimmer stimuli.

Previous research has failed to show a significant correlation between the PSSs from the SJ and TOJ tasks (e.g. Van Eijk *et al.*, 2008). The PSSs estimated from the current experiment are plotted in Fig. 5. While an effect of intensity is apparent across both tasks, the PSSs were uncorrelated at all three intensity levels (dim: *r*(16) = 0.22, *p* = 0.397; test: *r*(16) = 0.44, *p* = 0.077; bright: *r*(16) = 0.28, *p* = 0.272), confirming our data’s consistency with past findings.

1. *The Effect of Preceding-Intensity-Driven Perceived Offsets on Perceived Simultaneity*

The previous section shows a systematic effect of intensity on PSS and confirms that our behavioural data are consistent with past findings. The following section aims to assess the effect of the relative intensity and SOA of the preceding trial on the PSS. We therefore split our behavioural data, both by the visual intensity (*n* − 1 = dim or *n* − 1 = bright) and by the SOA (−200AV to 200VA) of the preceding trial. The data for each observer were individually plotted, based on the preceding intensity and SOA, to obtain the mean PSSs depicted in Fig. 6.

For the TOJ data, an effect of preceding SOA that is in line with the prediction (Fig. 2) was evident, with a larger magnitude shift for trials preceded by larger offsets. Here, a 25-ms difference was observed between the mean PSS of trials preceded by a 200-ms audio-leading stimulus, in comparison to a 200-ms visual-leading stimulus. Importantly, an effect of preceding intensity was also evident, with a shift that is larger than predicted. A regression line was plotted to the PSSs of the trials, dependent on the intensity of the preceding trial (dim versus bright). The difference between the intercepts of the regression line fitted to the trials preceded by dim (*MEAN =* −37.53ms, *SD* = 580) and bright (*MEAN =* −22.26ms, *SD* = 55.29) trials was significant (*t*(16) = 1.93, *p* = 0.036, *d* = 0.46) (one-tailed).

For the SJ data, the effect of preceding SOA was also apparent, with a 21-ms mean PSS difference between trials preceded by a 200 ms audio-leading trial, in comparison to a 200-ms visual-leading trial. However, the regression lines fitted to the PSSs, dependent on the intensity of the preceding trial, appear to be closely aligned. The difference between the intercepts for trials preceded by dim (*MEAN =* 56.93 ms, *SD* = 38.26) and bright (*MEAN =* 54.89 ms, *SD* =34.43) trials was non-significant (*t*(16) = −0.42, *p* = 0.342, *d* = −0.10) (one-tailed).

To summarise, the results suggest that the PSS rapidly recalibrates, based on the intensity of the preceding trial. This shift was in line with our prediction in the TOJ task, though there was no effect in the SJ task.

1. *Reaction Times*

Figure 7 (top row) displays median RTs, across the seven SOAs and three visual intensities, prior to splitting the test intensity data based on the properties of the preceding stimulus. A repeated-measures ANOVA revealed a significant main effect of SOA (*F*1.49,25.30 = 53.34, *p* < 0.001, *η*p*2* = 0.76), and of intensity (*F*1.10,18.74 = 5.13, *p* = 0.033, *η*p*2* = 0.23), on RTs of the unsplit data. The interaction between SOA and intensity was non-significant (*F*1.13,19.17 = 4.00, *p* = 0.056, *η*p*2* = 0.19).

The main aim of our RT analysis, however, was to assess whether the intensity or leading modality of the preceding trial affected the RT of the current trial at the test intensity (Fig. 7, bottom row), indicative of a preceding-intensity- and preceding-SOA-dependent effect on processing latencies. While this analysis revealed a significant effect of SOA (*F*1.09,18.58 = 11.77, *p* = 0.002, *η*p*2* = 0.41), there was no effect of preceding intensity (*F*1,17) = 0.66, *p* = 0.430, *η*p*2* = 0.04) and no effect of preceding leading modality (*F*1,17 = 1.38, *p* = 0.257, *η*p*2* = 0.08). Furthermore, no interactions were significant (all *p* > 0.1).

Assuming RTs are reflective of relative processing latencies (see Section 4.3), the data presented here suggest that the effect of the intensity or leading modality of the preceding trial on the PSS (in the TOJ task), is not caused by early latency shifts.

1. **Discussion**

The current paper aimed to investigate the hypothesis that the perception of simultaneity rapidly recalibrates to compensate for relative, intensity-dependent central arrival times of audio-visual signals. It was predicted that dim stimuli would be processed at longer latencies than bright stimuli, and therefore the perception of simultaneity would shift to compensate for the resultant difference in central arrival times. Experiment 1 partially supported this hypothesis as the PSS rapidly recalibrated based on the intensity of the preceding trial in the TOJ task, and the direction of this shift was in line with our prediction. However, there was no effect of preceding intensity in the SJ or simple RT task.

1. *The Role of Intensity-Dependent Processing Latencies in the Perception of Simultaneity*

Processing latencies estimated using response times are dependent on visual intensity (Pöppel *et al.*, 1990). Our RT data showed a similar intensity dependence (see Fig. 7), most notably for the visual-leading trials at larger SOAs, suggesting observers responded after detecting the first arriving signal (Raab, 1962). Likewise, the PSSs estimated using the SJ and TOJ tasks shifted towards the visual-leading SOAs with reducing visual intensity, consistent with previously reported data (Boenke *et al.*, 2009; Horsfall *et al.*, 2021a; Jaśkowski, 1992; Leone and Mccourt, 2015; Menendez and Lit, 1983; Roufs, 1963). These findings contribute to the argument that the intensity dependency of perceived simultaneity is driven by relative intensity-dependent arrival latencies at multisensory brain regions (Horsfall, 2021a; Leone and McCourt, 2015).

1. *The Task-Specific Effect of Preceding Intensity on Perceived Simultaneity*

The PSS has been shown to rapidly recalibrate due to the leading modality of an immediately preceding trial (Noel *et al.*, 2016; Roseboom, 2019; Simon *et al.*, 2017; Van der Burg *et al.*, 2013, 2015). While this effect is widely replicated (including in our data), audio-visual signals in the real world occur simultaneously at their source, and therefore the reason for such a mechanism to exist is not known. Here we propose that a central purpose of rapid recalibration is to correct for rapidly varying stimulus intensities, in this case luminance.

Since the perception of simultaneity shifts based on the physical offset of a stimulus, it was hypothesised that intensity-dependent latency delays would induce an equivalent shift in the PSS (see Fig. 2). For the TOJ data, a preceding-intensity-dependent PSS shift was observed, and critically, the direction of this shift was in line with our prediction. Van der Burg *et al.* (2018) reported that rapid recalibration is based solely on physical audio-visual offsets, rather than perceived offsets. Our behavioural findings are inconsistent with this, by showing that rapid recalibration is dependent on relative (intensity-dependent) *physiological* arrivallatencies.

In contrast to the TOJ data, no effect of preceding intensity was observed on the PSS when estimated using the SJ task. While no statistical analysis was used to compare the effect of preceding intensity between the SJ and TOJ tasks, as this was beyond the scope of the current paper, there are documented differences between the tasks (e.g. Linares and Holcombe, 2014; Yarrow *et al.*, 2011), including an inconsistency in the effect of preceding SOA on the PSS (e.g., Recio *et al.*, 2019; Roseboom, 2019).

While the reason for the inconsistency between the current data of the two tasks is uncertain, it has previously been argued that it is possible to detect more subtle changes in the PSS using the TOJ versus the SJ task, due to the higher cognitive demand required to make TOJs (Shore *et al.*, 2005). While the additional cognitive demand required by the TOJ task can be corroborated by subjective reports (Love *et al.*, 2013) and functional imaging data (Miyazaki *et al.*, 2016), the mechanism of how this increases the task’s sensitivity to small PSS changes is unclear. One tentative explanation is that more attentional resources are attributed to the audio-visual stimuli when making the more difficult TOJs (Volosin and Horváth, 2020), resulting in a greater detection of subtle differences in signal arrival times. This would be in line with evidence that increased attention enhances a participant’s sensitivity to temporal offsets in perceptual judgements of audio-visual stimuli (Donohue *et al.*, 2015).

1. *Rapid Recalibration to Preceding Visual Intensity is Not Evident in Simple Reaction Times*

While RTs reflect an interaction between processing latencies and a RT-specific decisional criterion (Cao et al., 2007; Watson, 1986), and therefore do not represent absolute processing latencies, they *have* been shown to be correlated with neural processing latencies measured using electroencephalography (EEG) (Hülsdünker et al., 2019; Schomaker, 2009). It was therefore assumed that if rapid recalibration was a low-level adaptation effect driven by changes in early processing latencies, we would expect to see an effect of the preceding trial’s characteristics on simple RTs, provided that this was not counteracted by an opposing shift of similar magnitude in the RT-specific decisional criterion.

We observed no effect of preceding leading modality or intensity on RTs, and so our findings are at odds with the possibility that the properties of a single preceding trial are compensated for via a change in relative processing latencies, such as what is observed following prolonged exposure to temporally offset bimodal stimuli (Di Luca *et al.*, 2009; Navarra *et al.*, 2009). As far as we are aware, ours is the first experiment to investigate the effect of rapid recalibration on simple RTs. These findings are perhaps unsurprising, as comparable EEG research by Simon *et al.* (2017) has revealed that preceding-SOA-dependent rapid recalibration is reflected in differences in late-evoked components associated with perceptual decision-making (Kelly and O’Connell, 2013; O’Connell *et al.*, 2012; Tagliabue *et al.*, 2019), as opposed to changes in early processing latencies.

1. **Conclusions**

The current experiments highlight that perceived simultaneity, when estimated using the TOJ task, rapidly recalibrates based on the visual intensity of the preceding trial. While this effect was small (< 20 ms), it was larger than the predicted preceding-intensity-dependent shift. Furthermore, there was no analogous effect on simple RTs, thus providing no evidence that rapid recalibration to the visual intensity of the preceding trial is driven by changes in early processing latencies. Based on our findings, it may be expected that the rapid recalibration mechanism could also partially correct for changes in distance between an audio-visual event and the observer, albeit only when the associated audio-visual signals reach the observer with a small temporal offset, since recalibration approaches zero for larger temporal offsets (see Van der Burg *et al.*, 2013). Since other cues such as stimulus motion (Wuerger *et al.*, 2010) and distance (Kopinska and Harris, 2004) are taken into account when making perceptual judgements, further research using more ecologically valid stimuli will be required to see if such an effect occurs in real-world settings.

**References**

Basharat, A., Adams, M. S., Staines, W. R. and Barnett-Cowan, M. (2018). Simultaneity and temporal order judgments are coded differently and change with age: an event-related potential study, *Front. Integr. Neurosci.* **12**, 15. DOI: 10.3389/fnint.2018.00015.

Binder, M. (2015). Neural correlates of audiovisual temporal processing — Comparison of temporal order and simultaneity judgments, *Neuroscience* **300**, 432–447. DOI: 10.1016/j.neuroscience.2015.05.011.

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.

Burr, D., Silva, O., Cicchini, G. M., Banks, M. S. and Morrone, M. C. (2009). Temporal mechanisms of multimodal binding, *Proc. Biol. Sci.* **276**, 1761–1769. DOI: 10.1098/rspb.2008.1899.

Cao, D., Zele, A. J. and Pokorny, J. (2007). Linking impulse response functions to reaction time: Rod and cone reaction time data and a computational model, *Vision Res.* **47**, 1060–1074. DOI: 10.1016/j.visres.2006.11.027.

Carrillo-de-la-Peña, M., Rodríguez Holguín, S., Corral, M. and Cadaveira, F. (1999). The effects of stimulus intensity and age on visual-evoked potentials (VEPs) in normal children, *Psychophysiology* **36**, 693–698. DOI: 10.1111/1469-8986.3660693.

Di Luca, M., Machulla, T.-K. and Ernst, M. O. (2009). Recalibration of multisensory simultaneity: Cross-modal transfer coincides with a change in perceptual latency, *J. Vision* **9**, 7. DOI: 10.1167/9.12.7.

Diederich, A. and Colonius, H. (2015). The time window of multisensory integration: Relating reaction times and judgments of temporal order, *Psychol. Rev.* **122**, 232–241. DOI: 10.1037/a0038696.ß

Donohue, S. E., Green, J. J. and Woldorff, M. G. (2015). The effects of attention on the temporal integration of multisensory stimuli, *Front. Integr. Neurosci.* **9**, 32. DOI: 10.3389/fnint.2015.00032.

Fujisaki, W., Shimojo, S., Kashino, M. and Nishida, S. (2004). Recalibration of audiovisual simultaneity, *Nat. Neurosci.* **7**, 773–778. DOI: 10.1038/nn1268

García-Pérez, M. A. and Alcalá-Quintana, R. (2012). On the discrepant results in synchrony judgment and temporal-order judgment tasks: a quantitative model, *Psychon. Bull, Rev.* **19**, 820–846. DOI: 10.3758/s13423-012-0278-y.

Gondan, M., Niederhaus, B., Rösler, F. and Röder, B. (2005). Multisensory processing in the redundant-target effect: A behavioral and event-related potential study, *Percept. Psychophys.* **67**, 713–726. DOI: 10.3758/BF03193527.

Horsfall, R., Wuerger, S. and Meyer, G. (2021a). Visual intensity-dependent response latencies predict perceived audio–visual simultaneity, *J. Math. Psychol.* **100**, 102471. DOI: 10.1016/j.jmp.2020.102471.

Horsfall, R. P., Wuerger, S. M. and Meyer, G. F. (2021b). Narrowing of the audiovisual temporal binding window due to perceptual training is specific to high visual intensity stimuli, *Iperception.* **12**. DOI: 10.1177/2041669520978670.

Hülsdünker, T., Ostermann, M. and Mierau, A. (2019). The speed of neural visual motion perception and processing determines the visuomotor reaction time of young elite table tennis athletes, *Front. Behav. Neurosci.* **13**, 165. DOI: 10.3389/fnbeh.2019.00165.

Jaskownki, P., Rybarczyk, K. and Jaroszyk, F. (1994). The relationship between latency of auditory evoked potentials, simple reaction time, and stimulus intensity, *Psychol. Res.* **56**, 59–65. DOI: 10.1007/BF00419712.

Jaśkowski, P. (1991). Two-stage model for order discrimination, *Percept. Psychophys.*, **50**, 76–82. DOI: 10.3758/BF03212206.

Jaśkowski, P. (1992). Temporal-order judgment and reaction time for short and long stimuli, *Psychol. Res.* **54**, 141–145. DOI: 10.1007/BF00922093.

Jaśkowski, P., Rybarczyk, K., Jaroszyk, F., & Lemański, D. (1995). The effect of stimulus intensity on force output in simple reaction time task in humans, *Acta Neurobiol. Exp.* **55**, 57–64. DOI: 10.55782/ane-1995-1061.

Kelly, S. P., & O’Connell, R. G. (2013). Internal and external influences on the rate of sensory evidence accumulation in the human brain, *J. Neurosci.* **33**, 19434–19441. DOI: 10.1523/JNEUROSCI.3355-13.2013.

Kopinska, A. and Harris, L. R. (2004). Simultaneity constancy, *Perception*, **33**, 1049–1060. DOI: 10.1068/p5169.

Leone, L. and McCourt, M. E. (2012). The question of simultaneity in audiovisual multisensory integration, *Proc. SPIE,* **8291**, 82910J1–J8. DOI: 10.1117/12.912183.

Leone, L. M. and McCourt, M. E. (2013). The roles of physical and physiological simultaneity in audiovisual multisensory facilitation, *iperception* **4**, 213–228, DOI: 10.1068/i0532.

Leone, L. M. and Mccourt, M. E. (2015). Dissociation of perception and action in audiovisual multisensory integration, *Eur. J. Neurosci.* **42**, 2915–2922. DOI: 10.1111/ejn.13087.

Linares, D. and Holcombe, A. O. (2014). Differences in perceptual latency estimated from judgments of temporal order, simultaneity and duration are inconsistent, *Iperception* **5**, 559–571. DOI: 10.1068/i0675.

Lines, C. R., Rugg, M. D. and Milner, A. D. (1984). The effect of stimulus intensity on visual evoked potential estimates of interhemispheric transmission time, *Exp. Brain Res.* **57**, 89–98. DOI: 10.1007/BF00231135.

Love, S. A., Petrini, K., Cheng, A. and Pollick, F. E. (2013). A psychophysical investigation of differences between synchrony and temporal order judgments, *PLoS One* **8**, e54798. DOI: 10.1371/journal.pone.0054798.

Menendez, A. and Lit, A. (1983). Effects of test-flash and steady background luminance on simple visual reaction time (RT) and perceived simultaneity (PS), *Invest. Ophthalmol. Vis. Sci.* **24**, 5.

Miyazaki, M., Kadota, H., Matsuzaki, K. S., Takeuchi, S., Sekiguchi, H., Aoyama, T. and Kochiyama, T. (2016). Dissociating the neural correlates of tactile temporal order and simultaneity judgements, *Sci. Rep.* **6**, 23323. DOI: 10.1038/srep23323.

Navarra, J., Hartcher-O’Brien, J., Piazza, E. and Spence, C. (2009). Adaptation to audiovisual asynchrony modulates the speeded detection of sound, *Proc. Natl Acad. Sci. U. S. A.*, **106**, 9169–9173. DOI: 10.1073/pnas.0810486106.

Noel, J. P., De Niear, M., Van der Burg, E. and Wallace, M. T. (2016). Audiovisual simultaneity judgment and rapid recalibration throughout the lifespan, *PLoS One* **11**, e0161698. DOI: 10.1371/journal.pone.0161698.

O’Connell, R. G., Dockree, P. M. and Kelly, S. P. (2012). A supramodal accumulation-to-bound signal that determines perceptual decisions in humans, *Nat. Neurosci.* **15**, 1729–1735. DOI: 10.1038/nn.3248.

Parise, C. V. and Ernst, M. O. (2016). Correlation detection as a general mechanism for multisensory integration, *Nat. Commun.* **7**, 11543. DOI: 10.1038/ncomms11543.

Pins, D. and Bonnet, C. (1997). Reaction times reveal the contribution of the different receptor components in luminance perception, *Psychon. Bull. Rev.* **4**, 359–366. DOI: 10.3758/BF03210793.

Plainis, S., Petratou, D., Giannakopoulou, T., Radhakrishnan, H., Pallikaris, I. G. and Charman, W. N. (2013). Interocular differences in visual latency induced by reduced-aperture monovision, *Ophthalmic Physiol. Opt.* **33**, 123–129. DOI: 10.1111/opo.12018

Pöppel, E., Schill, K. and von Steinbüchel, N. (1990). Sensory integration within temporally neutral systems states: A hypothesis, *Naturwissenschaften*, **77**, 89–91. DOI: 10.1007/BF01131783.

Raab, D. H. (1962). Statistical facilitation of simple reaction times, *Trans. N. Y. Acad. Sci.* **24**, 574–590. DOI: 10.1111/j.2164-0947.1962.tb01433.x.

Raij, T., Ahveninen, J., Lin, F.-H., Witzel, T., Jääskeläinen, I. P., Letham, B., Israeli, E., Sahyoun, C., Vasios, C., Stufflebeam, S., Hämäläinen, M. and Belliveau, J. W. (2010). Onset timing of cross-sensory activations and multisensory interactions in auditory and visual sensory cortices, *Eur. J. Neurosci.* **31**, 1772–1782. DOI: 10.1111/j.1460-9568.2010.07213.x.

Recio, R. S., Cravo, A. M., de Camargo, R. Y. and van Wassenhove, V. (2019). Dissociating the sequential dependency of subjective temporal order from subjective simultaneity, *PLoS One* **14**, e0223184. DOI: 10.1371/journal.pone.0223184.

Roseboom, W. (2019). Serial dependence in timing perception, *J. Exp. Psychol. Hum. Percept. Perform.* **45**, 100–110. DOI: 10.1037/xhp0000591.

Roufs, J. A. J. (1963). Perception lag as a function luminance, *Vision Res.* **3**, 81–91.

Schomaker, J. (2009). The relationship between response time and the strength of top-down attentional control: an ERP study, *J. Eur. Psychol. Stud.* **1**, 2. DOI: 10.5334/jeps.ab.

Shore, D. I., Gray, K., Spry, E. and Spence, C. (2005). Spatial modulation of tactile temporal-order judgments, *Perception*, **34**, 1251–1262. DOI: 10.1068/p3313.

Simon, D. M., Noel, J. P. and Wallace, M. T. (2017). Event related potentials index rapid recalibration to audiovisual temporal asynchrony, F*ront. Integr. Neurosci.* **11**, 8. DOI: 10.3389/fnint.2017.00008.

Simon, D. M., Nidiffer, A. R. and Wallace, M. T. (2018). Single trial plasticity in evidence accumulation underlies rapid recalibration to asynchronous audiovisual speech, *Sci. Rep.* **8**, 12499. DOI: 10.1038/s41598-018-30414-9.

Stelmach, L. B. and Herdman, C. M. (1991). Directed attention and perception of temporal order, *J. Exp. Psychol. Hum. Percept. Perform.* **17**, 539–550. DOI: 10.1037//0096-1523.17.2.539.

Symonds, M. R. E. and Moussalli, A. (2011). A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using Akaike’s information criterion, *Behav. Ecol. Sociobiol.* 65, 13–21. DOI: 10.1007/s00265-010-1037-6.

Tagliabue, C. F., Veniero, D., Benwell, C. S. Y., Cecere, R., Savazzi, S. and Thut, G. (2019). The EEG signature of sensory evidence accumulation during decision formation closely tracks subjective perceptual experience, *Sci. Rep.* **9**, 4949. DOI: 10.1038/s41598-019-41024-4.

Ulrich, R., Rinkenauer, G. and Miller, J. (1998). Effects of stimulus duration and intensity on simple reaction time and response force, *J. Exp. Psychol. Hum. Percept. Perform.* **24**, 915–928. DOI: 10.1037/0096-1523.24.3.915.

Van der Burg, E., Alais, D. and Cass, J. (2013). Rapid recalibration to audiovisual asynchrony, *J. Neurosci.* **33**, 14633–14637. DOI: 10.1523/JNEUROSCI.1182-13.2013.

Van der Burg, E., Orchard-Mills, E. and Alais, D. (2014). Rapid temporal recalibration is unique to audiovisual stimuli, *Exp. Brain Res.* **233**, 53–59. DOI: 10.1007/s00221-014-4085-8.

Van Der Burg, E., Alais, D. and Cass, J. (2015). Audiovisual temporal recalibration occurs independently at two different time scales, *Sci. Rep.* **5**, 14526. DOI: 10.1038/srep14526.

Van der Burg, E., Alais, D. and Cass, J. (2018). Rapid recalibration to audiovisual asynchrony follows the physical—not the perceived—temporal order, *Atten. Percept. Psychophys.* **80**, 2060–2068. DOI: 10.3758/s13414-018-1540-9.

Van Eijk, R. L. J., Kohlrausch, A., Juola, J. F. and van de Par, S. (2008). Audiovisual synchrony and temporal order judgments: Effects of experimental method and stimulus type, *Percept. Psychophys.* **70**, 955–968. DOI: 10.3758/PP.70.6.955.

Vatakis, A., Navarra, J., Soto-Faraco, S. and Spence, C. (2008). Audiovisual temporal adaptation of speech: temporal order versus simultaneity judgments, *Exp. Brain Res.* **185**, 521–529. DOI: 10.1007/s00221-007-1168-9.

Volosin, M. and Horváth, J. (2020). Task difficulty modulates voluntary attention allocation, but not distraction in an auditory distraction paradigm, *Brain Res.* **1727**, 146565. DOI: 10.1016/J.BRAINRES.2019.146565.

Watson, A. B. (1986). Temporal Sensitivity. In K. R. Boff, L. Kaufman and J. P. Thomas (Eds.). *Handbook of Perception and Human Performance*, Vol. 1 (pp. 6-1–6-43). Wiley, New York, NY, USA.

Watson, A. B. and Pelli, D. G. (1983). Quest: A Bayesian adaptive psychometric method, *Percept. Psychophys.,* **33**, 113–120. DOI: 10.3758/BF03202828.

Wuerger, S., Meyer, G., Hofbauer, M., Zetzsche, C. and Schill, K. (2010). Motion extrapolation of auditory–visual targets, *Inf. Fusion*, **11**, 45–50. DOI: 10.1016/J.INFFUS.2009.04.005.

Yarrow, K., Jahn, N., Durant, S. and Arnold, D. H. (2011). Shifts of criteria or neural timing? The assumptions underlying timing perception studies, C*onsc. Cogn.* **20**, 1518–1531. DOI: 10.1016/j.concog.2011.07.003.

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.

Zampini, M., Guest, S., Shore, D. I. and Spence, C. (2005). Audio-visual simultaneity judgments, *Percept. Psychophys.* **67**, 531–544. DOI: 10.3758/BF03193329.

**Appendix**

To determine how to best fit the data for the effect of preceding SOAs on the PSS (i.e. in Fig. 6), we compared the goodness of fit for two possible models. To do this, we compared two models that were fitted to the test intensity data (i.e., Fig. 6, dark blue circular markers). For Model 1, the regression lines (described in Section 2.8), comprising two parameters were fitted to the data:

 *a* + *bX* (A1)

 For model 2, a four-parameter sigmoidal curve (Boltzman) was fitted:

 *1*(*a − b*)/{1 + exp[(*x* *−* *c*)/*d*]} + *b* (A2)

where *a* represents the upper asymptote, *b* represents the lower asymptote, *c* reflects the location of the curve, and *d* reflects the width.

The corrected Akaike information criterion (AICc), which discourages overfitting by taking into account the number of parameters, was used to compare the models. This was calculated using the formula from Symonds and Moussalli (2011):

 AICc *= n*[ln(RSS/*n*)] + 2*k* + [2*k*(*k* + 1)/(*n* – *k* − 1)] (A3)

where *n* represents the number of data points, *k* represents the number of parameters, and RSS is the residual sum of squares, with lower values indicating a better fit.

The AICc for the two models was calculated for each observer separately, for both the SJ and TOJ data, and were then compared in two paired *t* tests. For the SJ data, the AICc values were significantly lower for Model 1 (MEAN *=* 43.70, SD = 6.38) than Model 2 (*MEAN =* 62.72, SD = 7.23) *t*16 = −43.52, *p* < 0.001, *d* = 10.56. Likewise, for the TOJ data, the AICc values were significantly lower for Model 1 (MEAN *=* 46.45, SD = 4.24) than Model 2 (MEAN *=* 65.54, SD = 4.00) *t*16 = −27.56, *p* < 0.001, *d* = 27.66, indicating that Model 1 best explained the current data for both tasks.

**Figure 1.** Depiction of an example trial order. Each trial consisted of an audio-visual (speaker and LED symbols) stimulus, that was either audio-leading, visual-leading or simultaneous. Trials were presented so that every other trial was at the ‘test’ intensity, and that these test intensity trials were always preceded by either a dim (grey), bright (cyan) or filler (see below) trial. Observers were asked to respond to every trial.

**Figure 2.** Prediction of how the point of subjective simultaneity (PSS) of the current trial (trial *N*) would shift to compensate for preceding audio-visual stimuli (trial *N* − 1) with different visual intensities. Panels A, B and C reflect schematic diagrams of how the perception of a single physical stimulus offset (A) can be affected by visual intensity-dependent processing latencies (B, C). In this example, an auditory beep is presented 50 ms before a visual flash (A). The bright visual stimulus (B) is processed at a shorter latency than the auditory stimulus, and therefore the perceived (audio-leading) offset between the audio and visual signals is reduced in comparison to the physical offset. The dim visual stimulus (C) has a longer processing latency than the auditory stimulus, and therefore the audio-leading offset is increased, relative to the physical offset. Panel D displays a quantitative prediction of a PSS shift, based on the preceding trial’s visual intensity and physical offset. The blue dashed line represents a regression line fitted to the predicted PSS, as a function of preceding stimulus onset asynchrony (SOA). The horizontal dotted blue line represents a mean predicted PSS for the test intensity stimulus, which was predicted by taking the average of the PSS at the 0.08 and 0.34 cd/m2 intensities of the temporal-order judgement (TOJ) data from Horsfall et al. (2021a). The magnitude of the preceding-intensity-dependent shift was predicted by mapping the physical temporal offset (A), alongside the intensity-dependent delay, onto the preceding-SOA-dependent shift (Simon et al., 2017). In the above example (A\, C), a −50-ms audio-leading trial, with a dim visual stimulus, would be perceived as a 75-ms audio-leading trial, and the perceptual system would recalibrate based on this perceived 75-ms offset. The predicted effect of preceding intensity is indicated by the cyan (N − 1 = bright) and grey (N − 1 = dim) dashed lines.

**Figure 3.** Intensity dependence of temporal order judgement (TOJ, left) and simultaneity judgement (SJ, right). Mean fits across observers in the TOJ (left) and SJ (right) tasks at the three visual intensities. Data points represent mean relative frequencies of ‘light first’ (temporal-order judgement, TOJ) or ‘simultaneous’ (SJ) responses (+/- SEM). Dashed lines represent the mean point of subjective simultaneity (PSS) at each intensity. The ‘test’ intensity data iare later split based on the properties of the preceding stimulus. The mean estimated PSSs are as follows: TOJ dim: 38 ms, TOJ test intensity: −29 ms, TOJ bright: −70 ms, SJ dim: 93 ms, SJ test intensity: 58 ms, SJ bright: 33 ms. SOA, stimulus onset asynchrony.

**Figure 4.** Individual fits for each observer (black dotted lines) in the temporal-order judgement (TOJ, top row) and simultaneity judgement (SJ, bottom row) tasks, at the dim (left column), test (middle column) and bright (right column) intensities. Each larger dot represents a relative frequency of ‘light first’ responses in the TOJ task, and ‘simultaneous’ responses in the SJ task. Thick lines represent mean fits. SOA, stimulus onset asynchrony.

**Figure 5.** Scatterplot of the points of subjective simultaneity (PSSs) recorded from the simultaneity judgement (SJ) and temporal-order judgement (TOJ) task. Circles represent the PSS for each individual participant at one of three intensities.

**Figure 6.** The effect of preceding intensity and stimulus onset asynchrony (SOA) on the point of subjective simultaneity (PSS) in the temporal-order judgement (TOJ, left) and simultaneity judgement (SJ, right) tasks. Regression lines were fitted to the PSSs of trials preceded by dim (solid grey line) and bright (solid cyan line) intensity stimuli, as a function of preceding SOA across all observers. The mean difference between the intercepts was 15 ms. The blue regression line reflects the medium-intensity data that were not split based on preceding intensity, and the blue dotted line reflects the mean of the PSSs at the test intensity. Data points reflect mean PSSs (+/− SEM). The predicted shifts were based on the current TOJ data (see Section 2.9 for how these predictions were made).

Figure 7 – Top row: Bimodal reaction times (RTs) at three visual intensities (dim – grey, test intensity – blue, and bright – cyan), across all stimulus onset asynchronies (SOAs). Bars represent the average RT, calculated across each observer’s median RT, for each SOA. Error bars represent standard deviations. Bottom row: Bimodal reaction times across all stimulus onset asynchronies (SOAs) at the test intensity, split based on the properties of the preceding stimulus. Bars represent the average RT across each observer’s median RT, dependent on whether the preceding stimulus was dim (grey bars), bright (cyan bars), auditory-leading (dashed outer line) or visual-leading (solid outer line). Our analysis revealed an effect of SOA, but no effect of the intensity or leading modality of the preceding trial on RTs.