CrossMark

RESEARCH ARTICLE

Multisensory integration is independent of perceived simultaneity

Vanessa Harrar[1,2] · Laurence R. Harris[3] · Charles Spence[2]

Abstract The importance of multisensory integration for perception and action has long been recognised. Integrating information from individual senses increases the chance of survival by reducing the variability in the incoming signals, thus allowing us to respond more rapidly. Reaction times (RTs) are fastest when the components of the multisensory signals are simultaneous. This response facilitation is traditionally attributed to multisensory integration. However, it is unclear if facilitation of RTs occurs when stimuli are perceived as synchronous or are actually physically synchronous. Repeated exposure to audiovisual asynchrony can change the delay at which multisensory stimuli are perceived as simultaneous, thus changing the delay at which the stimuli are integrated—perceptually. Here we set out to determine how such changes in multisensory integration for perception affect our ability to respond to multisensory events. If stimuli perceived as simultaneous were reacted to most rapidly, it would suggest a common system for multisensory integration for perception and action. If not, it would suggest separate systems. We measured RTs to auditory, visual, and audiovisual stimuli following exposure to audiovisual asynchrony. Exposure affected the variability of the unisensory RT distributions; in particular, the slowest RTs were either speed up or slowed down (in the direction predicted from shifts in perceived simultaneity). Additionally, the multisensory facilitation of RTs (beyond statistical summation) only occurred when audiovisual onsets were physically synchronous, rather than when they appeared simultaneous. We conclude that the perception of synchrony is therefore independent of multisensory integration and suggest a division between multisensory processes that are fast (automatic and unaffected by temporal adaptation) and those that are slow (perceptually driven and adaptable).

Introduction

Environmental events typically generate multiple types of energy with profound differences in the speed of propagation, the time taken for the senses to transduce that energy, and the time taken by the resulting signals to travel through the central nervous system. For example, a person speaking at a distance will generate sounds that will reach the ear after the corresponding visual signals have reached the eye. Since it is highly unlikely that signals corresponding to simultaneous auditory and visual stimuli occurring in the natural world arrive in the brain simultaneously (except at a distance of 10 m or so, the so-called horizon of simultaneity; Pöppel et al. 1990; Spence and Squire 2003), it has been proposed that the brain monitors the temporal correspondence between multisensory signals and adjusts for regular asynchronies (Kopinska and Harris 2004). The

✉ Vanessa Harrar
vanessa.harrar@umontreal.ca

1 School of Optometry, University of Montreal, 3744 Jean-Brillant, Montreal, QC H3T 1P1, Canada

2 Department of Experimental Psychology, University of Oxford, Oxford OX1 3UD, UK

3 Centre for Vision Research, York University, Toronto, ON M3J 1P3, Canada

temporal delay can, within limits, be adjusted when determining the true temporal relationship between the signals, as long as the distance from the event to the observer is known (Kopinska and Harris 2004).

In support of this "temporal adjustment theory", the perception of multisensory synchrony has been found to be plastic (Fujisaki et al. 2004; Vroomen et al. 2004). Following repeated exposure to an asynchronous stimulus, the point of subjective simultaneity (PSS) shifts towards the repeatedly experienced delay (for a review see Vroomen and Keetels 2010). Anecdotally, this phenomenon can be experienced when watching a movie with an out-of-synch soundtrack: after a few minutes, the auditory delay becomes less perceptible and less annoying. It is generally assumed that perceived stimulus timing reflects the arrival time of the independent signals at some critical place in the brain (Gibbon and Rutschmann 1969; Harrar and Harris 2005; Ulrich 1987). Thus, the common explanation for changes in perceptual simultaneity following temporal adaptation is that the processing speed of the independent signals has changed, causing the resultant changes in perceived simultaneity (Navarra et al. 2009; though see Harrar and Harris 2008). What is not known or understood is how this flexibility in perceived simultaneity is related to the neural process of multisensory integration. Like the perception of simultaneity, multisensory integration also requires signals to occur close together in time, within a so-called integration window (Stein and Stanford 2008).

The benefits of multisensory integration are closely related to the temporal relationship between the stimuli. Physically simultaneous multisensory stimuli give rise to faster and more accurate performance (Ernst and Banks 2002; Stein and Meredith 1993; Van der Burg et al. 2008). These advantages are generally lost when pairs of stimuli are separated by a temporal interval that is too large, particularly if they are perceived as arising from separate events (Colonius and Diederich 2004; Meredith et al. 1987). Thus, temporal adaptation, which shifts the temporal requirements for the perception of simultaneity, might also shift the integration window, and hence the temporal relationships associated with the largest multisensory behavioural benefits. To test this idea, we measured reaction times (RTs) to unisensory and multisensory stimulus pairs with various physical asynchronies between the component stimuli after exposure to different temporal adaptation conditions.

Post-adaptation unisensory mean RTs were compared across conditions. This traditional analysis, however, ignores the fact that RTs are not normally distributed, and instead have a long (positive) tail. We therefore also fitted the RT data with an ex-Gaussian distribution (Luce 1986), the convolution of a Gaussian function and an exponential function. The ex-Gaussian has three parameters: the mean

and standard deviation of the Gaussian component (μ and σ) and the rate of exponential component (τ). While the Gaussian component represents the bulk of the responses, the exponential component characterises the slower RTs that are indicative of the involvement of slower analytic and demanding processes (Balota and Spieler 1999). The more skewed the distribution of the data, relative to a normal distribution, the larger τ will be. Importantly, the sum of μ and τ equals the traditionally calculated mean of the RTs, and so by presenting both sets of unisensory RT summary statistics, we can replicate and extend previous research. Thus, using the ex-Gaussian fits, we can determine whether the previously reported differences in mean RT across adaptation conditions are attributable to a change in the Gaussian component (μ), a change in the exponential component (τ), or to a change in both.

Post-adaptation multisensory RTs were compared against the predictions of Miller's race model (Miller 1982). The race model describes the advantages in response speed predicted by the probability summation of two independent signals rather than a single signal (i.e. no special integration mechanisms are required). The race model states $F_{VA} \leq F_V(t) + F_A(t)$, where F_V and F_A are cumulative density functions (CDFs) of the two unisensory RTs (visual and auditory), while F_{VA} is the CDF of multisensory RTs in the audio–visual stimulus condition (see MATLAB algorithm in Ulrich et al. 2007). The race model prediction for multisensory RTs is the sum of the visual and auditory RT probability distributions and is therefore also known as "statistical summation" or "probability summation". The race model has been used to explain the temporal window of integration (Colonius and Diederich 2004, 2006). However, multisensory RTs are often even faster than the predictions of the simple race model (Molholm et al. 2002). These violations of the race model, RTs faster than predicted from statistical summation, indicate an additional advantage-producing process—the "magic" of multisensory integration. Do violations of the race model only occur when stimuli are within a temporal window centred on physical simultaneity, or might multisensory behavioural benefits shift with the ephemeral "perception of simultaneity"? This is the first time that RTs to multisensory stimulus combinations have been evaluated after temporal adaptation and the consequent shift in the PSS.

Participants were exposed to three adaptation regimes: audio-lagging, audio-leading, or synchronous stimulation. Following exposure, half of the participants then performed a temporal order judgement (TOJ) *perceptual* task where they reported whether the visual or auditory component of an audiovisual stimulus pair had been presented first. Methods and details of the TOJ results are presented in the online supplementary materials—results are as expected from the previous literature. The remainder of the

participants performed a *motor response* task where they pressed a button as soon as any stimulus was presented (the same button for visual, auditory, or multisensory stimuli). The pattern of response times, as a function of temporal adaptation, and stimulus onset asynchrony, are presented in the current manuscript.

Methods

Participants

Sixteen participants (10 female; median age 23 years) completed the study. The participants were randomly assigned to the audio-lagging or visual-lagging condition (8 participants in each). Participants from both groups were also tested in the synchronous-adaptation condition, either the week before or the week after the asynchronous condition. For each condition, testing was spread over 6 sessions, completed within one week in order to have enough RTs at each of the SOAs (100), with top-up adaptation periods after every four test-trials. It was important to have enough data at each SOA to generate representative cumulative distribution functions (where medians are calculated at every tenth percentile), so that the race model could be tested at each SOA. Participants therefore completed a total of twelve sessions, which lasted 35 min each. One participant did not complete the experiment; their partial data were not included in the analysis that follows. The participants gave their informed consent, and the University Research Ethics Committee at Oxford University approved the study.

Stimuli

Adapting stimuli were loud and high contrast to maintain attention during the no-response portion of the study. During testing, the stimuli were weaker and harder to detect in order to maximise integration effects (see inverse effectiveness in Stein and Meredith 1993). The auditory stimulus consisted of a burst of white noise (70 dB for adaptation; 45 dB for testing) played from two speakers placed on either side of the screen at the same height as the visual stimulus so that the audio and visual stimuli were co-localised. The visual stimulus consisted of a vertical Gabor patch (2° diameter, 0.3 cycles per degree, 0.822 Michelson contrast for adaptation; 0.05 contrast for testing) presented on a CRT computer monitor set to an 85 Hz refresh rate. The duration of all stimuli was 20 ms (with some variability for the visual stimulus duration due to refresh rate).

The multisensory stimuli presented during the test phase were different on each trial. The test stimuli were separated by SOAs of ± 250, ± 100, ± 80, ± 60, ± 40, or 0 ms. The synchrony of the auditory and visual stimuli was verified with an oscilloscope, where a photodiode used to determine that the onset of the visual stimulus was the same as the auditory stimulus. The multisensory stimuli were interleaved with unisensory visual and auditory stimuli. RTs were measured from the onset of the first signal. Stimuli were presented, and keyboard responses were recorded with E-Prime 2.0.

There were three adaptation stimuli: audio leading by 60 ms, visual leading by 200 ms, or synchronous. The asymmetrical adaptation delays were chosen to match the asymmetrical integration window; there is a higher tolerance for visual-leading as compared to audio-leading delays (Dixon and Spitz 1980). Regardless of the adaptation condition, the test stimuli were the same for each participant in each condition.

Procedure

Participants alternated between testing and adaptation blocks during a 5 min practice period before the experiment began. Immediately after practice, the participants began an initial 5 min adaptation period, where they were encouraged to count the audiovisual stimulus pairs in order to maintain attention on the stimuli and to encourage them to interpret each stimulus pair as a single unit. Immediately after the adaptation period, the test phase began with a short instruction displayed on the screen: "Testing. Respond after each trial". The participants pressed a button (a single button regardless of the stimulus) as rapidly as possible whenever they perceived a visual, auditory, or audiovisual stimulus (the stimuli were intermingled). The next trial was initiated after a random delay of 750–1250 ms after the response, to reduce predictability. There were six top-up adaptation trials after every four test-trials preceded by the message: "Adapting, pay attention". In total, there were 70 top-up blocks within each session, giving rise to a total of 420 top-up adaptation trials in each session, dispersed between 290 RT trials to unisensory and multisensory stimuli. After the six sessions with a given adaptation condition, we had collected 100 RTs at each of the 11 SOAs (total of 1100 multisensory RTs) and nearly 300 RTs for each unisensory stimulus, for each participant. Another six sessions were repeated the following week in a different adaptation condition.

Data analysis

All of the data presented here were from testing sessions only; no data were collected during the adaptation phase. Comparisons across conditions are only made between responses collected during the test portion, where participants were exposed to identical stimuli regardless of the condition. Summary statistics for unisensory RTs were

handled in two very different ways. First, as in previous papers in this field, we removed outliers (values greater than 500 ms post-stimulus, or ± 3 standard deviations: as in Di Luca et al. 2009; Harrar and Harris 2008; Navarra et al. 2009) and calculated a mean for each participant and then used statistics to determine any differences due to adaptation condition. However, several researchers have demonstrated that mean RT is not sufficient and it is important to look at additional distributional properties of RTs (even as far back as Donders 1868; Woodworth 1938). Ratcliff and Murdock (1976) points out that when the mean RT is speeded up, it is unclear if the overall RTs decreased, or if only the slower responses were speeded up. Ratcliff demonstrated that the convolution of a normal distribution with an exponential distribution provides a good summary of the shape of RT data and provides parameter values that are easy to interpret. Therefore, we complemented the description of mean unisensory RT with parameters of the ex-Gaussian distribution, using the functions in the MAT-LAB toolbox DISTRIB (Lacouture and Cousineau 2008). Since the ex-Gaussian distribution allows for a distribution with a tail, we did not need to crop the data. This analysis took into account all the data (including the few exceedingly late RT initially discredited as "outliers").

A generalised estimating equation (GEE, see Hanley et al. 2003) was used to determine any significant effects of temporal adaptation on unisensory RTs (same model for all dependent variables: mean, and parameters from the ex-Gaussian distribution μ, σ, and τ). In the GEE model, there were three within-subject factors (modality [audio or visual], session [1–6], and adaptation condition [AV, synchronous, or VA]) and an exchangeable correlation matrix structure (standard in the SPSS 16 library). Significant interaction effects, when present, were followed up with appropriate pairwise comparisons. In order to maintain a reasonably low alpha-rate, pairwise comparisons were only conducted within a modality (the obvious differences between modalities are not novel and not important). Thus, we followed up interactions by testing the effect of adaptation on RT for each modality individually.

Results

Results 1: Unisensory reaction time

Using standard RT analysis, in particular so that data here can be compared with previously published data, RTs of less than 100 ms or greater than 500 ms were removed, and the mean RT was calculated for each participant, in each condition (synchronous and asynchronous). The GEE analysis revealed that there was a significant interaction between the modality of the stimulus that the participants

were responding to and the adaptation condition (Wald $\chi^2 = 11.25$, $p = .004$). Based on previous results, we expected a progressive effect of adaptation delay on RT, but in the opposite directions for auditory and visual RTs. Indeed, the significant interaction indicates that the increasing slope of auditory RTs relative to the exposure delays is significantly different from the decreasing slope of visual RTs relative to the exposure delay (see Fig. 1a). The pairwise comparison was not significant for auditory RTs ($p = .16$); auditory RTs were 10 ms (± 7; 95% confidence interval [CI] −24 to +4 ms) faster following exposure to audio-lagging stimulus pairs compared to audio-leading pairs. In contrast, the pairwise comparison was significant for visual RTs ($p = .001$); visual RTs were 24 ms (±7; 95% CI 10 to 38 ms) faster following exposure to audio-leading compared to audio-lagging pairs.

The RT data were then analysed by fitting an ex-Gaussian distribution and comparing the parameters of the best-fit model for each participant across conditions and modality (group data and fits plotted in Fig. 1b). Ex-Gaussians were equally well fit to the data from all three adaptation conditions and modalities (r^2 values reported in Table 1). Using the GEE as described above, we observed a significant interaction (difference in slopes) between adaptation condition and modality for all three parameters: the mean/peak of the Gaussian (μ): Wald $\chi^2 = 9.006$, $p = .011$; the standard deviation of the Gaussian (σ): Wald $\chi^2 = 19.173$, $p < .001$; and the rate of the exponential component (τ): Wald $\chi^2 = 11.667$, $p = .003$), see means in Table 1. The interaction effect indicates that temporal adaptation affects visual and auditory RTs differently—indeed, inspection of the means in Table 1 reveals the pattern to be exactly opposite (decreasing for visual but increasing for auditory RTs). Follow-up analysis revealed that the effect of adaptation was not robust enough to produce significant pairwise differences of μ for either auditory or visual RTs.

In contrast, the effect of adaptation resulted in several significant differences in the variability measures for both auditory and visual RTs. Sigma, the standard deviation of the Gaussian component, was significant for both visual RTs [adapt visual-lead was larger than adapt to visual-lag (mean difference = 10.9 ms, SE = 3.4, $p = .001$) and adapt synchronous was larger than visual-lag (mean difference = 7.8 ms, SE = 2.0, $p < .001$)] and auditory RTs [adapt audio–lead was larger than audio–lag (mean difference = 7.1 ms, SE = 3.2, $p = .026$) and adapt audio–lead was greater than synchronous (mean difference = 5.1 ms, SE = 2.0, $p = .012$)]. Tau, the rate of the exponential component, varied significantly with conditions for visual RTs, in the same pattern as for sigma [adapt visual–lead was larger than visual–lag (mean difference = 22.9 ms, SE = 10.9, $p = .036$) and adapt to synchronous was larger than adapt to visual-lag (mean difference = 16.6 ms, SE = 8.1, $p = .040$)].
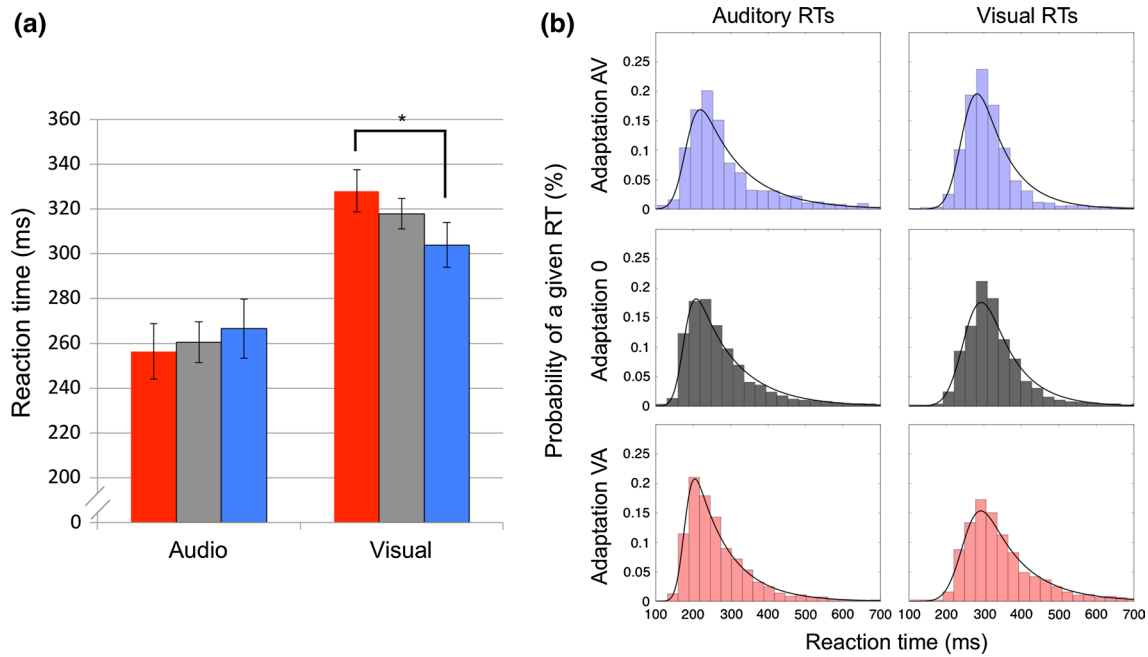
Fig. 1 RTs following adaptation. In *red* are results following adaptation to an audio-visual stimulus where auditory lags; in *blue* are results following adaptation to an audio-visual stimulus where the auditory leads; in *grey* are results following adaptation to a synchronously presented audio-visual stimulus. **a** Mean unisensory RTs following each of three adaptation conditions, using standard outlier removal processes. The *error bars* indicate the within-group standard errors calculated from mean RTs across all 12 participants. **b** Histogram and ex-Gaussian functions fit to reaction times in response to unisensory audio and visual stimuli (colour figure online)

Table 1 Parameters from ex-Gaussian fits to unisensory RT after three temporal adaptation procedures

Modality	Adapt condition	Mu Mean of Gaussian		Sigma Variability of Gaussian		Tau Exponential component		GoF
		Mean	SE	Mean	SE	Mean	SE	R^2
Visual	Visual-lead	259	9	38	4	92	12	.98
	Synchronous	260	6	35	2	76	8	.97
	Visual-lag	255	9	27	1	69	9	.96
Audio	Visual-lead	188	11	16	3	80	9	.99
	Synchronous	189	7	19	1	94	8	.98
	Visual-lag	202	10	24	2	90	11	.95

As a measure of goodness of fit (GoF), we also include the r^2 for each ex-Gaussian fit

Thus, the initially reported changes in the mean RTs (see Fig. 1a) are unlikely to be caused by changes in the whole distribution of responses. Instead, the changes in the mean RT are due mostly to changes in the variability and in the exponential component of the reaction time distribution (i.e. the slower responses were speeded or slowed down).

Results 2: Multisensory RTs

Multisensory RTs showed a similar pattern across adaptation conditions. Median post-adaptation RTs for visual, auditory, and audiovisual stimuli are plotted in Fig. 2a–c

for the three adaptation conditions as a function of the SOA. Note that the lines representing unisensory RTs differ slightly from one figure to the next, corresponding to the changes in unisensory RTs in response to temporal adaptation. The complete set of multisensory RTs for synchronous audio–visual stimuli (SOA = 0) are plotted in histograms in Fig. 2d, with the best-fit ex-Gaussian plotted through the data, demonstrating that adaptation also changed the multisensory RT. In order to determine whether the change in multisensory RTs is novel, or simply attributable to the changes in unisensory RTs, further processing was necessary.
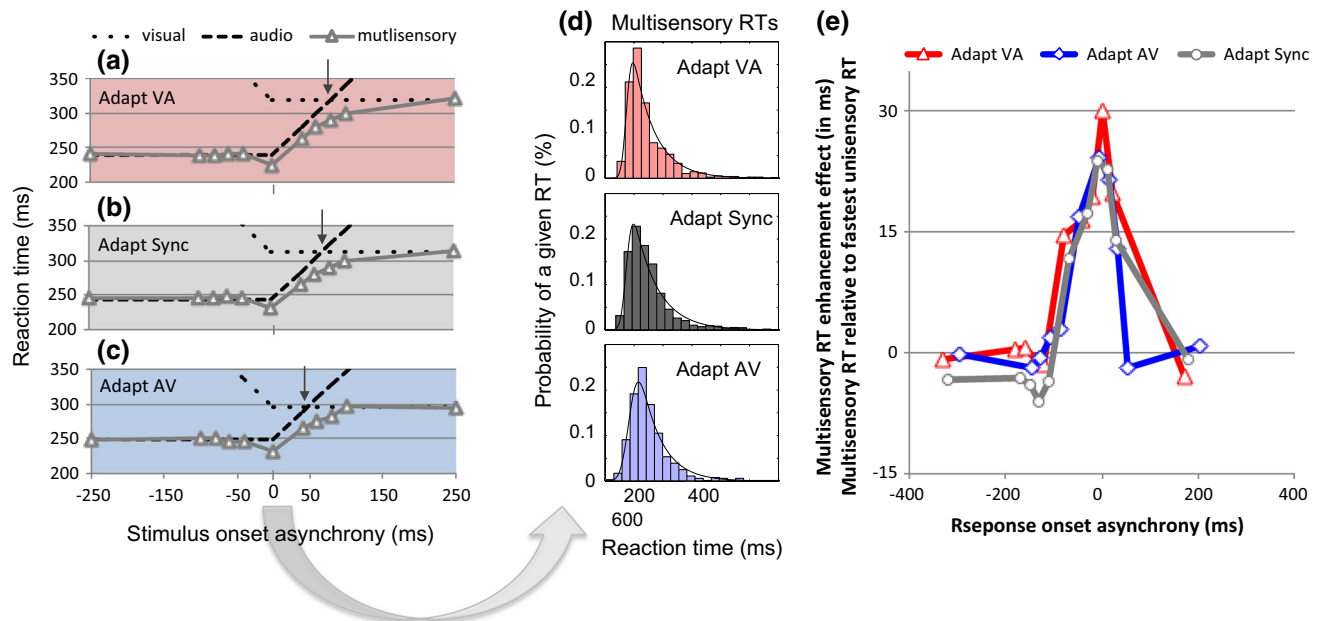
Fig. 2 Statistical facilitation of multisensory reaction times (RTs) relative to unisensory RTs. **a-c** Plots median RTs to auditory (*dashed*) and visual (*dotted lines*), and audio-visual stimuli (*triangles*) at each SOA. Unisensory RTs take into account the manually imposed delays (negative SOA means that the auditory stimulus was presented first). **a** *Red*, adapt to audio-visual stimulus in which audio is lagging. **b** *Grey*, adapt to synchronously presented audio-visual stimuli. **c** *Blue*, adapt to audio-visual stimulus in which audio is leading. While the unisensory RTs vary across adaptation conditions, multisensory RTs were always fastest when the stimuli were simultaneous (SOA = 0 ms). For this delay, we have plotted the entire distribu-

tion of multisensory RTs (depicted by a large sweeping *arrow* from SOA = 0 to histograms in d). **d** Histogram and ex-Gaussian functions fit to RTs in response to physically synchronous (SOA = 0) multisensory stimuli. **e** The total facilitation effect: differences between multisensory RTs and the fastest of the unisensory RTs are plotted against the response onset asynchrony (ROA), which is difference between unisensory RTs. When the ROA was zero (indicated by *arrows* pointing down in **a-c**), there is the largest statistical facilitation effect— yielding multisensory RTs that are ~20 ms faster than the fastest unisensory RT (colour figure online)

To calculate multisensory enhancement, we subtracted the multisensory RT from the fastest unisensory RT (for each individual, for each condition). Note that all participants were faster at responding to auditory-only stimuli than to visual-only stimuli. The absolute difference between the unisensory RTs was used to calculate each individual's response onset asynchrony (ROA). The ROA is the hypothesised difference between unisensory RTs at a given SOA. For example, the small black arrows in Fig. 2a–c depict the ROA of zero—where RTs are predicted to be the same for both unisensory signals; left of the arrows auditory RTs are faster, and right of the arrows visual RTs are faster. Thus, ROA = 0 indicates the SOA is such that the auditory stimulus is delayed by the appropriate amount so that response times to the auditory and visual stimuli should be equal. The ROA was calculated for each individual, in each condition. The mean enhancement effect is plotted relative to the ROA in Fig. 2e. Figure 2e demonstrates that for all three adaptation conditions, the largest multisensory RT enhancement was found when the ROA equals zero, when the visual stimulus leads the auditory stimulus by the difference in their unisensory RTs.

Results 3: Race model facilitation and violation

We then used the unisensory RTs (the actual raw data, not the summary statistics provided in Results 1) to generate a prediction for the multisensory RTs for each individual, in each of the adaptation conditions, at each response onset asynchrony, based on the principles of statistical summation. A certain amount of the multisensory RT facilitation is predicted by statistical summation of the unisensory signals. We used the race model detailed in Ulrich (Ulrich et al. 2007) importantly adding the SOA to the unisensory RTs so that the unisensory RTs were delayed by the appropriate amount to predict the multisensory RTs at each SOA. As a first step to determining the race model, cumulative density functions were generated for the unisensory and multisensory RTs. That is, raw RTs (for each individual in a given condition) were ordered and then divided in order to estimate ten percentiles (details in Ulrich et al. 2007). Inspection of RTs at each percentile (for each SOA) revealed that only when the multisensory stimuli were simultaneous, the multisensory RTs were faster than the race model. This increase was only true at the earliest
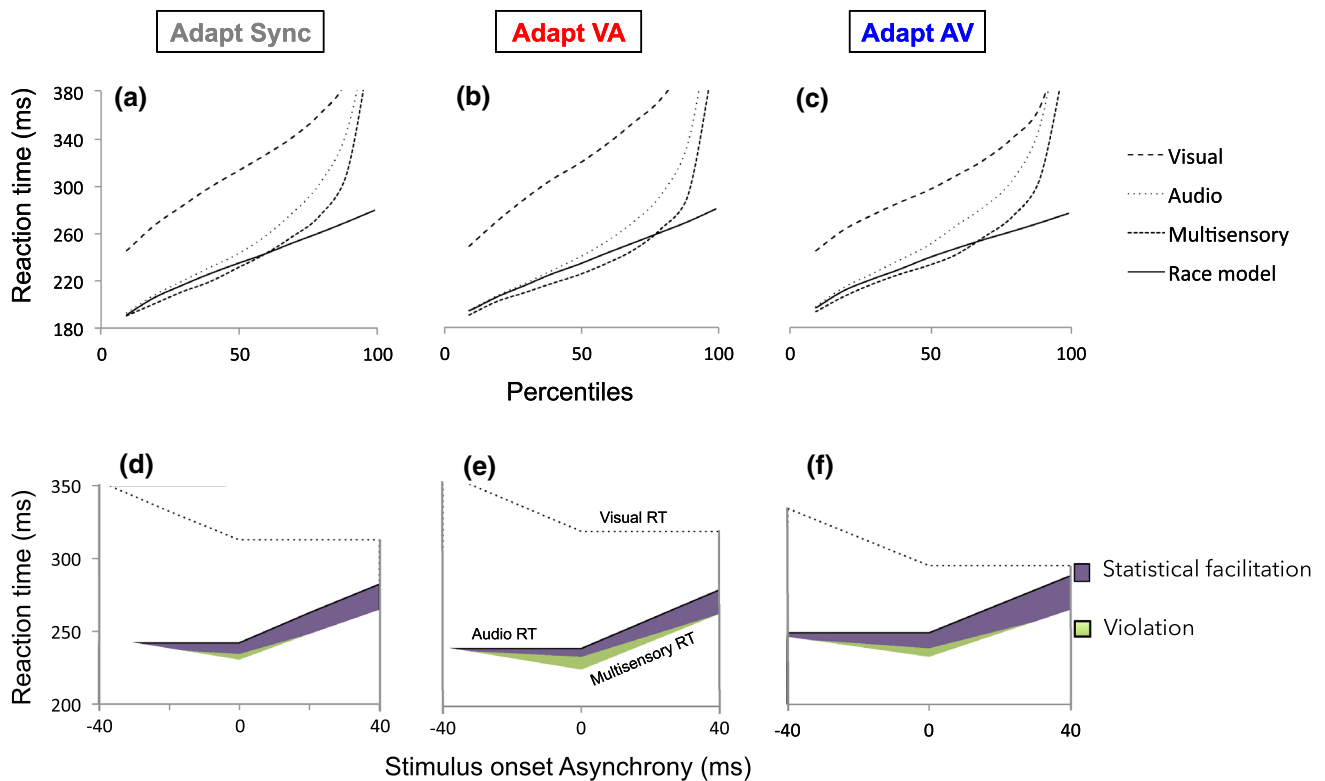
Fig. 3 Multisensory RTs can be modelled as a combination of statistical facilitation and "true" multisensory facilitation, in the three adaptation conditions. **a-c** cumulative density functions (CDFs) of the visual (*long dash*), auditory (*dotted*), multisensory (*dash*) RTs are plotted, as well as the race model (*solid line*), for each temporal adaptation condition. **d-f** Plots the RTs as function of three SOAs: ± 40, and 0. Statistical facilitation is demonstrated by the race model having a lower RT than the unisensory RTs (*purple region in* **d-f**, *solid line* lower than unisensory lines in **a-c**), while multisensory facilitation is observed when the multisensory RTs are faster than the race model (*green region in* **d-f**, multisensory line lower than *solid line* in **a-c**) (colour figure online)

percentiles: a distinctive sign of multisensory integration. This can be seen in Fig. 3a–c, which plots the cumulative density functions for the three temporal adaptation conditions, in the specific situation where the stimuli were presented simultaneously (SOA = 0).

The changes in the unisensory RT (caused by the adaptation paradigm, as reported in Results 1: Unisensory) produce different Race models for each adaptation condition. By comparing the multisensory RTs to the race model, we can determine how much of the multisensory enhancement is due to statistical facilitation, or multisensory facilitation. Figure 3d–f demonstrates that the enhancement of RTs is made up of two parts: statistical facilitation (predicted by the race model, in purple) and multisensory facilitation (violations of the race model, in green). These violations of the race model indicate an additional facilitation of RTs that cannot be explained by the statistical summation of the unisensory signals. We have termed this additional facilitation: multisensory facilitation.

In order to calculate the magnitude of multisensory facilitation at each SOA, we subtracted the multisensory RTs from the predictions of the race model; negative values thus indicate violations of the race model. Figure 4a plots the magnitude of multisensory facilitation relative to the SOA for each adaptation condition. If temporal adaptation affected multisensory processes, for example by changing the correlation between the unisensory signals, then we would expect the pattern of violation to vary across adaptation condition (exactly predicted are illustrated by the displaced vertical arrows in Fig. 4a). In fact, the data collected following our three temporal asynchrony regimes demonstrated nearly identical violation patterns. To compare the violations at each SOA, across the three adaptation conditions, we analysed the data with a GEE: SOA (11) × Condition (3) × week[1] (2). Testing the simple effects of adapt condition at each SOA (in order to determine whether the violations at each SOA are different between groups), revealed no significant effect of temporal adaptation at any

---

[1] The GEE is similar to the one described in "Data analysis" with a few minor differences. The within-subjects variables were SOA and "week" since in order to have enough data at a given SOA, multisensory RT data needed to be combined across sessions.
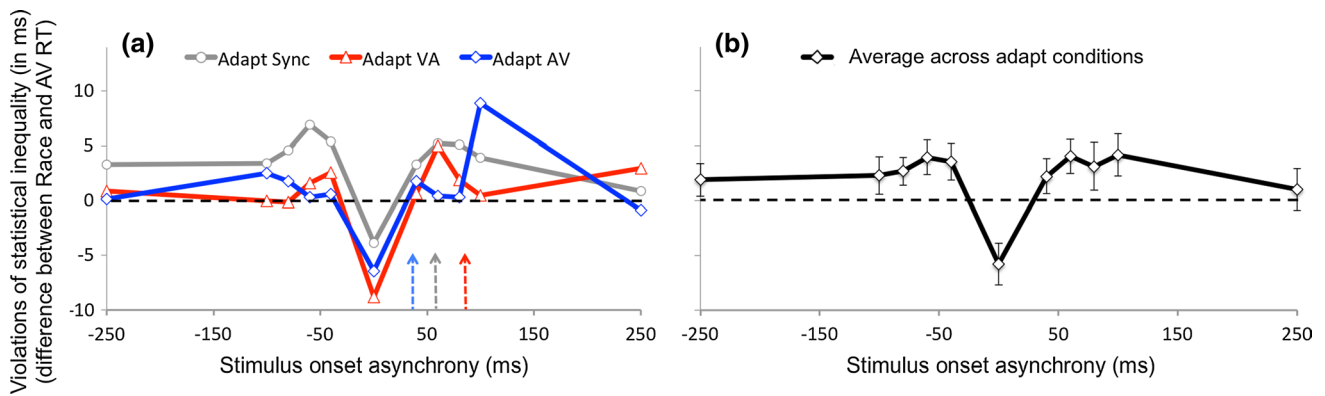
Fig. 4 Race model violations. The violations of the race model are plotted as a function of the SOA between the components of the multisensory stimuli. Violations below zero (*horizontal dashed line*) correspond to RTs that were faster than the values predicted by statistical summation. Negative SOAs correspond to audio first. **a** *Red lines* indicate violations in the condition where participants adapted to audio--lagging; *blue lines* indicate adaptation to audio--leading; *grey lines* indicate adaptation to synchronous. *Dotted vertical arrows* pointing up indicate the SOAs at which the stimuli were approximately "physiologically simultaneous" (based on RTs) for each adaptation condition. Violations indicating multisensory facilitation were only found when stimuli were physically synchronous—i.e. the same violation pattern is observed regardless of the temporal adaptation condition. **b** The average data, combined across the three adaptation conditions, *error bars* indicate the standard error of the mean (colour figure online)

of the 11 SOAs (Wald χ values ranged from 0.265 to 5.869, corresponding to *p* values ranging from .876 to .053). This confirmed that the pattern of race-model violations was not reliably affected by the adaptation procedure even though the unisensory RTs changed.

The amount of violation did, however, vary as a function of the SOA; there was a significant main effect of SOA (Wald $\chi_{10} = 191.8$, $p < .001$). Pairwise comparisons revealed that the violation values were, for the most part, not different across SOAs ($p > .05$ for 90 pairwise comparisons), except at SOA = 0—which was significantly different from all other violation values ($p < .01$ comparing violation at SOA = 0 to all ten other SOAs). Next, we wanted to know whether these violations were significantly different from zero—significantly greater than zero (slower) would indicate inhibition of responses, while significantly less than zero (faster) would indicate response facilitation. We followed up with Bonferroni-corrected one-sample *t* tests at each SOA, comparing the RT violations to zero. The only temporal relationship that resulted in RTs that were significantly different from the race model prediction was when the multisensory stimuli were simultaneous (SOA = 0). When the stimuli were simultaneous, RTs were faster than predicted by the race model; the mean violation was 5.8 ms (SE = 1.9 ms, $t_{29} = -3.02$, $p = .005$, see Fig. 4b).

In summary, the small differences between violation patterns across the three temporal adaptation conditions were not reliable. Overall, the three temporal adaptation conditions produced the same violation pattern at each SOAs. Importantly, only when the stimuli were simultaneous (SOA = 0) was there reliable evidence multisensory facilitation of RTs. There was some indication that at other SOAs, when either visual or audio led, there might have been a response inhibition where RTs were slower than predicted by the model.[2]

General discussion

Temporal adaptation to audiovisual asynchrony changed the mean unisensory RTs to both component stimuli, in the direction predicted by the delay used during the adaptation phase (replicating Di Luca et al. 2009; Harrar and Harris 2008; Navarra et al. 2009). Analysing the RT by fitting an ex-Gaussian distribution revealed that the aforementioned changes in the mean RT can mostly be attributed to changes in the variability of the RTs (σ) and a change in the late exponential component (τ). Multisensory RTs were analysed by comparing them to the race model. We found that the pattern of multisensory RTs relative to the statistical summation model was the same across adaptation procedures. Multisensory facilitation (response enhancement beyond statistical facilitation) only occurred for physically simultaneous stimuli (SOA = 0). Since temporal

---

[2] Using an uncorrected alpha-rate, there were several additional SOAs that produced RTs that were significantly *slower* than the race model demonstrating a trend towards response inhibition: SOA = -80 ms, mean response inhibition ($\bar{x}$) = 2.6 ms ± (SE)1.7 ms, $t_{29} = 2.09$, $p = .045$; SOA = -60 ms, $\bar{x} = 3.9 \pm 1.6$ ms, $t_{29} = 2.52$, $p = .018$; SOA = -40 ms, $\bar{x} = 3.5 \pm 1.6$ ms, $t_{29} = 2.14$, $p = .041$; SOA = 60 ms, $\bar{x} = 4.0 \pm 1.6$ ms, $t_{29} = 2.59$, $p = .015$; SOA = 100 ms, $\bar{x} = 4.1 \pm 1.9$ ms, $t_{29} = 2.18$, $p = .038$.

adaptation is known to change multisensory perception of synchrony, but does not appear to affect multisensory response enhancement, we suggest that multisensory action and perception might be based on different components of the neural response (early or late, respectively).

Multisensory RTs were in line with the previous reports in two important ways. First, statistical facilitation (defined as the increased speed of responding to multisensory stimuli relative to unisensory stimuli) was always largest when one stimulus was delayed by the SOA that would make it physiologically synchronous with the other stimulus in the brain (response onset asynchrony: ROA). In the data presented here, the largest statistical facilitation occurred when the slower (visual) stimulus was given a head-start so that the two signals arrived at the brain at the same moment (or at least were responded to with the same speed). Statistical facilitation is purely driven by the variability in the responses and is therefore completely attributed to the probability summation of two stimuli (i.e. the race model, Diederich and Colonius 2004; Hershenson 1962; Miller 1986) and is not necessarily related to the multisensory nature of the stimulus. The principles of variability and congruent effectiveness explain the types of benefits observed in multisensory behaviours (Otto et al. 2013).

Second, the fastest absolute RT was observed when the stimuli were presented simultaneously. The race model also predicts this. However, in addition, multisensory facilitation, measured by violations of the race model, only occurred when the signals were presented synchronous at source, i.e. physically synchronous (see also Leone and McCourt 2012, 2013) regardless of the temporal adaptation paradigm. Using adaptation, which we know is associated with changes in multisensory synchrony perception, we were able to change the unisensory RTs and demonstrate that multisensory facilitation of responses remained stable. Multisensory facilitation of RTs is, thus, independent of factors that affect processing speed and perception.

Multisensory RTs faster than the race model

Race model violations have previously been reported for physically synchronous multisensory stimuli (see earliest reports in Miller 1986; Molholm et al. 2002). There have been several suggestions as to what might cause multisensory response facilitation beyond statistical summation. For example, the "modality shift effect" postulated to cause delays in unisensory RTs when the stimuli are not blocked by modality (Spence et al. 2001a), certainly accounts for a small portion of the multisensory facilitation (Miller 1986). More recently, models that allow the correlation between the unisensory signals to vary can also account for the multisensory facilitation (Otto et al. 2013), as well

as models that allow reduced variability in the neuronal response of the unisensory stimuli (Otto and Mamassian 2012). The results of the current analysis, demonstrating the effect of temporal adaptation on the variability and the exponential component of the unisensory RTs, support Otto and Mamassian's hypothesis that the variability in unisensory RTs depends on multisensory experience. Otto and Mamassian's revisions generalise the race model making it more difficult to be falsified (violated), because in their model the unisensory responses are allowed to vary depending on the context. Our data reveal that these kinds of context-dependent unisensory responses are only (if ever) necessary when multisensory stimuli have a physically simultaneous onset—even when a nonzero delay between the stimuli is more likely to be perceived as simultaneous (as in the current data set). Indeed, none of the revisions to the race model, or explanations for multisensory facilitation, have considered the relevance of perception to behavioural responses. This is the first report where multisensory RTs at various SOAs were tested after the multisensory temporal structure was perceptually altered. While unisensory response profiles changed in the direction predicted by perceptual changes, multisensory facilitation did not. We found that the largest violation of the race model consistently occurred at "true simultaneity" (when the stimuli were physically simultaneous at source), even in conditions in which the point of perceptual synchrony would have been shifted. In addition to violations of the race model demonstrating ultra-fast responses, there were also a number of SOAs where inhibition might have occurred—where median RTs appear to be slower than predicted by the race model (see Fig. 4b).

Multisensory RTs slower than the race model

Inhibition of multisensory RTs has also been seen before (Leone and McCourt 2013; Miller 1986), and at similar SOAs as the current study (e.g. in Leone and McCourt 2012), but has yet to be adequately explained. The same type of inhibition is also reported physiologically; single-cell recordings from a subset of multisensory neurons (cells that fire to multisensory stimuli and when either auditory or visual stimuli are presented alone) show response inhibition (less firing to multisensory stimuli than to either unisensory stimulus). The maximum level of response depression has been reported as occurring when the visual stimulus was initiated 100 ms before the auditory stimulus (see Fig. 10 in Meredith et al. 1987). A recent publication suggests an excellent model in which sub-additive responses in multisensory neurons are the result of an interaction of inhibitory and excitatory inputs of multisensory cells, which can result in a depressed firing rate in response

to a multisensory stimulus, relative to the response of the unisensory stimuli (Miller et al. 2015). Extrapolating from the response patterns of such cells, we can suggests some kind of inter-sensory inhibitory pathway that might cause these slower RTs reported here at SOAs 50–100 ms. Further research is needed to identify under what conditions inhibition and facilitation of responses can be modelled from the multisensory combination of unisensory responses, especially in light of the independence between the behavioural and perceptual effects.

Unisensory RT changes can at least partially explain shifts in the perception of synchrony following adaptation (Fujisaki et al. 2004; Vroomen et al. 2004). After adaptation to visual-leading pairs (VA), the mean response times to flashes are slower (Fig. 1a, also Di Luca et al. 2009; Harrar and Harris 2008; Navarra et al. 2009), and RTs to flashes are associated with a larger exponential component (i.e. more late RTs, see Fig. 1b; Table 1). Thus, some neurological process appears to slow down the internal visual signal (maybe by holding it in a buffer, as suggested by Jiang et al. 1994), and/or speed-up the internal auditory signal, to compensate for the offset and create a synchronised perception. Such a calibration system is impressively specific; changes in synchrony perception of one person speaking do not generalise to another person speaking, demonstrating an important contextual element (Roseboom and Arnold 2011). The context of speech appears to cause important differences in the size of the integration window for the auditory and the visual components of speech, sometimes making it smaller (e.g. Maier et al. 2011) and sometimes making it larger (e.g. Wallace and Stevenson, 2014). An alternative to the buffering model is suggested as a result of the ex-Gaussian analysis of unisensory RTs, which clarified that temporal adaptation affects the variability of the RTs and the exponential component.

The two components of the ex-Gaussian RT distribution are likely related to two different underlying processes. The Gaussian portion of the distribution is thought to arise from early automatic (nonanalytic) processes. On the other hand, the exponential component can be an indication of energetical factors such as arousal (Sanders 1983), or the involvement of analytic and demanding processes that are slower (e.g. attention, Balota and Spieler 1999). Thus, the changes in the proportion of responses in the tail of the distribution, a change in the tau value, argue against a direct computation factor affecting RTs—such an effect would have caused the whole distribution of RTs to shift. Instead, temporal adaptation caused a change in the variability between individual trials. Adaptation appears to affect the allocation of energy resources, which are needed to maintain a consistent state of preparation for stimulus processing and response generation (Leth-Steensen et al. 2000). The first stimulus in the adaptation pair (the stimulus that leads) is characterised by a RT with a larger variability, indicated by the RTs having a larger tau, suggesting greater attentional lapses, engagement of additional slow (demanding) processes, and perhaps greater difficulty allocating effort to the first stimulus in the pair. On the other hand, the stimulus that lags in the adaptation pair appears to be characterised by RTs with a smaller exponential component, and a smaller variability, suggesting it engages fewer late cognitively demanding processes and recruits a steady level of effort across trials.

Changes in the exponential component of the unisensory RT distribution did not translate into changes in the pattern of multisensory RT facilitation, which suggests that multisensory violations of the race model are driven by the earlier (faster) Gaussian component of the RT distribution. Multisensory RTs are always faster than unisensory RTs, and are primarily driven by the fastest unisensory responses (as described in the race model). Thus, factors that change the exponential component of the unisensory RT distribution—characterised by late responses, either speeding them up or slowing them down—are unlikely to affect the multisensory RTs. Along these lines, perception, being a relatively slow process, is likely affected by the same factors that contribute to the later component of the RT distribution. Thus, temporal adaptation affects participants' arousal states towards individual stimuli, which causes changes to the exponential portion of the RT distribution, and simultaneously affects the perception of simultaneity (PSS), but does not affect the way that the unisensory stimuli come together to generate the super-fast multisensory facilitated responses. While there is some disagreement about interpreting the components of the ex-Gaussian in terms of cognitive processes (e.g. see Matzke and Wagenmakers, 2009), there seems to be considerable evidence here for dissociation between the RTs to multisensory stimuli, and the perceptions of such stimuli. We therefore propose for division between multisensory processes that are "slow" (contextual, analytic, criterion-driven, generally perceptual) such as synchrony perception and the McGurk effect (both of which are affected by temporal adaptation, Yuan et al. 2014), and multisensory integration processes that are "fast" (automatic, generally actions) such as reacting to a stimulus with finger press or eye movements (Colonius and Diederich 2004). Several hypotheses have been put forward suggesting the processes that underlie multisensory temporal perception, which we suggest would be unrelated to multisensory facilitation of RTs.

Sternberg and Knoll (1973) suggested that the perception of temporal order depends on arrival latencies of stimuli at a central comparator. Our ex-Gaussian analysis of unisensory RTs suggests that temporal adaptation affects attention, and attention to a sensory modality has been shown to affect multisensory arrival latencies and

synchrony perception by means of prior entry (Spence et al. 2001b). Prior entry is caused by a reduction in encoding time of the attended stimulus and a prolongation of the unattended stimulus (Vibell et al. 2007; Tünnermann et al. 2015). Temporal adaptation may cause the second stimulus in the adaptation pair (the one that lags) to be cued for the next stimulus presentation—causing it to be processed more quickly—which should cause the adaptation pair to appear more synchronous, and could also cause the lagging stimulus to be processed more quickly even when tested alone. These changes to the unisensory RT distribution after asynchrony adaptation may therefore be related to changes in the allocation of attention, which might underlie changes in the perception of synchrony. However, the asymmetric changes in temporal perception following temporal adaptation cannot be fully explained by changes in unisensory processing (see data and explanation in Roach et al. 2011).

Changes in the temporal perception (i.e. PSS) might also arise from changes in the gain of neurons selectively tuned to particular audiovisual asynchronies (Roach et al. 2011), such as the multisensory neurons in the anterior ectosylvian cortex (Berman 1961; Jiang et al. 1994). These neurons might also be able to deduce the distance of a multisensory event to the observer based on the asynchrony between the auditory and visual signals (Jaekl et al. 2012, 2015). While there is no solid evidence that temporal adaptation causes changes in neural timings (Roseboom et al. 2015; Yarrow et al. 2011), we propose that any such effects would be most pronounced in the slower responses of these neurons (cf. tau) and that it is these late responses which appear to be related to the perception of synchrony. These neuronal changes cannot, however, without further development, explain the dissociation between multisensory temporal perception and actions demonstrated here.

Alternatively, a Bayesian framework can account for flexible synchrony perception with a fixed temporal delay where multisensory RT facilitation occurs. In the Bayesian structure that might underlie multisensory temporal perception, there could be a "prior" (for actual synchrony) and a "likelihood" function that is adaptable (Hanson et al. 2008; Keetels and Vroomen 2012; Miyazaki et al. 2006). While the deeply ingrained prior would always assumes that events occurring close together in time and space are actually simultaneous, the likelihood function would be built-up over the course of a few minutes of experience (e.g. temporal adaptation) causing a shift to what the observer interprets as "close together in time". While changes to the likelihood function might result in changes to the PSS, the inflexibility of the prior may underlie the fixed nature of the multisensory response facilitation that only occurs for truly simultaneous signals—ones that correspond to the original prior.

Perception and action are often independent. In vision, they are classically represented by distinct neural networks (Goodale and Milner 1992). This dissociation also seems relevant in the multisensory research framework. Colonius and Diederich's (2004) time window of integration (TWIN) model demonstrated that the integration window is significantly larger for the RT task than for the TOJ task (Diederich and Colonius 2015). The dissociation between the RT and TOJ has been demonstrated by others and is proposed to be due to the fact that the information used in the two tasks is triggered from different processes with different critical decision thresholds (Jaskowski 1992). Our data support the distinction between multisensory integration for perception and integration for action. In the context of our data, processes that affect integration for perception (e.g. temporal adaptation affecting the PSS) appear to be independent from processes that affect the integration for multisensory facilitation of responses (see also Megevand et al. 2013). The data presented here further suggest that the "temporal window" for multisensory perception of synchrony is adjustable by experience, whereas the time window for fast multisensory action-based benefits is not.

Knowledge that multisensory behavioural benefits and multisensory perception are independent could perhaps be used to inform training programs for clinical populations with multisensory temporal deficits. We suggest that temporal adaptation training could speed unisensory processing; for example, watching a movie in which the audio track is delayed may speed subsequent auditory processing. Slower than normal unisensory processing appears to underpin a multisensory binding deficit found in patients with schizophrenia (Foucher et al. 2007) and dyslexics suffering from delayed auditory processing (Farmer and Klein 1995) causing multisensory deficits (Harrar et al. 2014). However, in those cases where unisensory processing appears to have developed normally and deficits only emerge in multisensory integration tasks (e.g. in some cases of autism spectrum disorder, see Foss-Feig et al. 2010), temporal adaptation is unlikely to affect their ability to process multisensory stimuli.

Conclusions

The dissociation presented here between changes in multisensory perception related to changes in unisensory processing, and multisensory interactions that are unchanged by temporal adaptation, is novel within the multisensory literature. The data presented here strongly support separate systems for the perception of a multisensory event and our ability to act on that event (Leone and McCourt 2015). Having used the exact same stimuli in the different conditions, varying only the adaptation asynchrony (which is

known to consistently affect perception), we provide evidence for the dissociation in multisensory integration for perception and action. As a tentative explanation for the necessity of this dissociation, we note that optimal performance in a TOJ task (perception) requires maximum segregation, while optimal performance in a multisensory RT task (action) requires maximum integration. A further novel insight, related to the shape of the distribution of RTs, was that multisensory action and perception might be based on different components of the neural response (early or late, respectively). The multisensory system for action is likely optimised to escape from predators as quickly as possible (an early basic process), which is quite different from demands of the multisensory perceptual system (which relies on slower analytic processes). Thus, the independence of integration for perception and integration for action might be driven by the difference in task demands (Megevand et al. 2013).

References

Balota DA, Spieler DH (1999) Word frequency, repetition, and lexicality effects in word recognition tasks: beyond measures of central tendency. J Exp Psychol Gen 128(1):32

Berman AL (1961) Interaction of cortical responses to somatic and auditory stimuli in anterior ectosylvian gyrus of cat. J Neurophysiol 24(6):608–620

Colonius H, Diederich A (2004) Multisensory interaction in saccadic reaction time: a time-window-of-integration model. J Cogn Neurosci 16(6):1000–1009. doi:10.1162/0898929041502733

Colonius H, Diederich A (2006) The race model inequality: interpreting a geometric measure of the amount of violation. Psychol Rev 113(1):148–154. doi:10.1037/0033-295X.113.1.148

Di Luca M, Machulla TK, Ernst MO (2009) Recalibration of multisensory simultaneity: cross-modal transfer coincides with a change in perceptual latency. J Vis 9(12):7.1–716. doi:10.1167/9.12.7

Diederich A, Colonius H (2004) Bimodal and trimodal multisensory enhancement: effects of stimulus onset and intensity on reaction time. Percept Psychophys 66(8):1388–1404

Diederich A, Colonius H (2015) The time window of multisensory integration: relating reaction times and judgments of temporal order. Psychol Rev 122(2):232–241. doi:10.1037/a0038696

Dixon NF, Spitz L (1980) The detection of auditory visual desynchrony. Perception 9(6):719–721

Donders FC (1868) /1969). Die Schnelligkeit psychischer Processe [On the speed of mental processes]. Acta Psychol 30:412–431

Ernst MO, Banks MS (2002) Humans integrate visual and haptic information in a statistically optimal fashion. Nature 415(6870):429–433. doi:10.1038/415429a

Farmer ME, Klein RM (1995) The evidence for a temporal processing deficit linked to dyslexia: a review. Psychon Bull Rev 2(4):460–493. doi:10.3758/Bf03210983

Foss-Feig JH, Kwakye LD, Cascio CJ, Burnette CP, Kadivar H, Stone WL, Wallace MT (2010) An extended multisensory temporal binding window in autism spectrum disorders. Exp Brain Res 203(2):381–389. doi:10.1007/s00221-010-2240-4

Foucher JR, Lacambre M, Pham BT, Giersch A, Elliott MA (2007) Low time resolution in schizophrenia Lengthened windows of simultaneity for visual, auditory and bimodal stimuli. Schizophr Res 97(1–3):118–127. doi:10.1016/j.schres.2007.08.013

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

Gibbon J, Rutschmann R (1969) Temporal order judgement and reaction time. Science 165(3891):413–415

Goodale MA, Milner AD (1992) Separate visual pathways for perception and action. Trends Neurosci 15(1):20–25

Hanley JA, Negassa A, Forrester JE (2003) Statistical analysis of correlated data using generalized estimating equations: an orientation. Am J Epidemiol 157(4):364–375

Hanson JV, Heron J, Whitaker D (2008) Recalibration of perceived time across sensory modalities. Exp Brain Res 185(2):347–352. doi:10.1007/s00221-008-1282-3

Harrar V, Harris LR (2005) Simultaneity constancy: detecting events with touch and vision. Exp Brain Res 166(3–4):465–473. doi:10.1007/s00221-005-2386-7

Harrar V, Harris LR (2008) The effect of exposure to asynchronous audio, visual, and tactile stimulus combinations on the perception of simultaneity. Exp Brain Res 186(4):517–524. doi:10.1007/s00221-007-1253-0

Harrar V, Tammam J, Perez-Bellido A, Pitt A, Stein J, Spence C (2014) Multisensory integration and attention in developmental dyslexia. Curr Biol 24(5):531–535. doi:10.1016/j.cub.2014.01.029

Hershenson M (1962) Reaction time as a measure of intersensory facilitation. J Exp Psychol 63:289–293

Jaekl P, Soto-Faraco S, Harris LR (2012) Perceived size change induced by audiovisual temporal delays. Exp Brain Res 216(3):457–462. doi:10.1007/s00221-011-2948-9

Jaekl P, Seidlitz J, Harris LR, Tadin D (2015) Audiovisual delay as a novel cue to visual distance. PLoS ONE 10(10):e0141125. doi:10.1371/journal.pone.0141125

Jaskowski P (1992) Temporal-order judgment and reaction time for short and long stimuli. Psychol Res 54(3):141–145

Jiang H, Lepore F, Ptito M, Guillemot J-P (1994) Sensory interactions in the anterior ectosylvian cortex of cats. Exp Brain Res 101(3):385–396

Keetels M, Vroomen J (2012) Exposure to delayed visual feedback of the hand changes motor-sensory synchrony perception. Exp Brain Res 219(4):431–440. doi:10.1007/s00221-012-3081-0

Kopinska A, Harris LR (2004) Simultaneity constancy. Perception 33(9):1049–1060

Lacouture Y, Cousineau D (2008) How to use MATLAB to fit the ex-Gaussian and other probability functions to a distribution of response times. Tutor Quant Methods Psychol 4(1):35–45

Leone LM, McCourt ME (2012) The question of simultaneity in multisensory integration. Human Vis Electron Imaging Xvii. doi:10.1117/12.912183

Leone LM, McCourt ME (2013) The roles of physical and physiological simultaneity in audiovisual multisensory facilitation. i-Perception 4(4):213–228. doi:10.1068/i0532

Leone LM, McCourt ME (2015) Dissociation of perception and action in audiovisual multisensory integration. Eur J Neurosci 42(11):2915–2922

Leth-Steensen C, Elbaz ZK, Douglas VI (2000) Mean response times, variability, and skew in the responding of ADHD children: a response time distributional approach. Acta Psychol (Amst) 104(2):167–190

Luce RD (1986) Response times. Oxford University Press, New York

Maier JX, Di Luca M, Noppeney U (2011) Audiovisual asynchrony detection in human speech. J Exp Psychol Hum Percept Perform 37(1):245

Matzke D, Wagenmakers EJ (2009) Psychological interpretation of the ex-Gaussian and shifted Wald parameters: a diffusion model analysis. Psychon Bull Rev 16(5):798–817. doi:10.3758/PBR.16.5.798

Megevand P, Molholm S, Nayak A, Foxe JJ (2013) Recalibration of the multisensory temporal window of integration results from changing task demands. PLoS ONE 8(8):e71608. doi:10.1371/journal.pone.0071608

Meredith MA, Nemitz JW, Stein BE (1987) Determinants of multisensory integration in superior colliculus neurons. I. Temp Factors. J Neurosci 7(10):3215–3229

Miller J (1982) Divided attention: evidence for coactivation with redundant signals. Cogn Psychol 14(2):247–279

Miller J (1986) Timecourse of coactivation in bimodal divided attention. Percept Psychophys 40(5):331–343

Miller RL, Pluta SR, Stein BE, Rowland BA (2015) Relative unisensory strength and timing predict their multisensory product. J Neurosci 35(13):5213–5220. doi:10.1523/JNEUROSCI.4771-14.2015

Miyazaki M, Yamamoto S, Uchida S, Kitazawa S (2006) Bayesian calibration of simultaneity in tactile temporal order judgment. Nat Neurosci 9(7):875–877. doi:10.1038/nn1712

Molholm S, Ritter W, Murray MM, Javitt DC, Schroeder CE, Foxe JJ (2002) Multisensory auditory–visual interactions during early sensory processing in humans: a high-density electrical mapping study. Brain Res Cogn Brain Res 14(1):115–128

Navarra J, Hartcher-O'Brien J, Piazza E, Spence C (2009) Adaptation to audiovisual asynchrony modulates the speeded detection of sound. Proc Natl Acad Sci U S A 106(23):9169–9173. doi:10.1073/pnas.0810486106

Otto TU, Mamassian P (2012) Noise and correlations in parallel perceptual decision making. Curr Biol 22(15):1391–1396. doi:10.1016/j.cub.2012.05.031

Otto TU, Dassy B, Mamassian P (2013) Principles of multisensory behavior. J Neurosci 33(17):7463–7474. doi:10.1523/JNEUROSCI.4678-12.2013

Pöppel E, Schill K, von Steinbuchel N (1990) Sensory integration within temporally neutral systems states: a hypothesis. Naturwissenschaften 77(2):89–91

Ratcliff R, Murdock BB (1976) Retrieval processes in recognition memory. Psychol Rev 83(3):190

Roach NW, Heron J, Whitaker D, McGraw PV (2011) Asynchrony adaptation reveals neural population code for audio–visual timing. Proc Biol Sci 278(1710):1314–1322. doi:10.1098/rspb.2010.1737

Roseboom W, Arnold DH (2011) Twice upon a time multiple concurrent temporal recalibrations of audiovisual speech. Psychol Sci 22(7):872–877

Roseboom W, Linares D, Nishida S (2015) Sensory adaptation for timing perception. Proc R Soc Lon B: Biol Sci 282:20142833. doi:10.1098/rspb.2014.2833

Sanders A (1983) Towards a model of stress and human performance. Acta Psychol (Amst) 53(1):61–97

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

Spence C, Nicholls ME, Driver J (2001a) The cost of expecting events in the wrong sensory modality. Percept Psychophys 63(2):330–336

Spence C, Shore DI, Klein RM (2001b) Multisensory prior entry. J Exp Psychol-Gen 130(4):799–832. doi:10.1037//0096-3445.130.4.799

Stein BE, Meredith MA (1993) The merging of the senses. MIT Press, Cambridge, Mass

Stein BE, Stanford TR (2008) Multisensory integration: current issues from the perspective of the single neuron. Nat Rev Neurosci 9(4):255–266. doi:10.1038/nrn2331

Sternberg S, Knoll RL (1973) The perception of temporal order: Fundamental issues and a general model. Attention and Performance IV:629–685

Tünnermann J, Petersen A, Scharlau I (2015) Does attention speed up processing? Decreases and increases of processing rates in visual prior entry. J Vis 15(3):1–27

Ulrich R (1987) Threshold models of temporal-order judgments evaluated by a ternary response task. Percept Psychophys 42(3):224–239

Ulrich R, Miller J, Schroter H (2007) Testing the race model inequality: an algorithm and computer programs. Behav Res Methods 39(2):291–302

Van der Burg E, Olivers CN, Bronkhorst AW, Theeuwes J (2008) Pip and pop: nonspatial auditory signals improve spatial visual search. J Exp Psychol Hum Percept Perform 34(5):1053–1065. doi:10.1037/0096-1523.34.5.1053

Vibell J, Klinge C, Zampini M, Spence C, Nobre AC (2007) Temporal order is coded temporally in the brain: early ERP latency shifts underlying prior entry in a crossmodal temporal order judgment task. J Cogn Neurosci 19:109–120

Vroomen J, Keetels M (2010) Perception of intersensory synchrony: a tutorial review. Atten Percept Psychophys 72(4):871–884. doi:10.3758/APP.72.4.871

Vroomen J, Keetels M, de Gelder B, Bertelson P (2004) Recalibration of temporal order perception by exposure to audio-visual asynchrony. Brain Res Cogn Brain Res 22(1):32–35. doi:10.1016/j.cogbrainres.2004.07.003

Wallace MT, Stevenson RA (2014) The construct of the multisensory temporal binding window and its dysregulation in developmental disabilities. Neuropsychologia 64:105–123. doi:10.1016/j.neuropsychologia.2014.08.005

Woodworth RS, Schlosberg H (1938) Experimental psychology. Holt, New York, pp 327–328

Yarrow K, Jahn N, Durant S, Arnold DH (2011) Shifts of criteria or neural timing? The assumptions underlying timing perception studies. Conscious Cogn 20(4):1518–1531

Yuan X, Bi C, Yin H, Li B, Huang X (2014) The recalibration patterns of perceptual synchrony and multisensory integration after exposure to asynchronous speech. Neurosci Lett 569:148–152