

# Operating in a Multisensory Context: Assessing the Interplay Between Multisensory Reaction Time Facilitation and Inter-sensory Task-switching Effects

Luke H. Shaw,<sup>a</sup> Edward G. Freedman,<sup>a</sup> Michael J. Crosse,<sup>b</sup> Eric Nicholas,<sup>a</sup> Allen M. Chen,<sup>a</sup> Matthew S. Braiman,<sup>a</sup> Sophie Molholm<sup>a,b</sup> and John J. Foxe<sup>a,b,\*</sup>

<sup>a</sup> The Cognitive Neurophysiology Laboratory, The Del Monte Institute for Neuroscience, Department of Neuroscience, University of Rochester School of Medicine and Dentistry, Rochester, NY 14642, USA

<sup>b</sup> The Cognitive Neurophysiology Laboratory, Department of Pediatrics & Neuroscience, Albert Einstein College of Medicine & Montefiore Medical Center, Bronx, NY 10461, USA

**Abstract**—Individuals respond faster to presentations of bisensory stimuli (e.g. audio-visual targets) than to presentations of either unisensory constituent in isolation (i.e. to the auditory-alone or visual-alone components of an audio-visual stimulus). This well-established multisensory speeding effect, termed the redundant signals effect (RSE), is not predicted by simple linear summation of the unisensory response time probability distributions. Rather, the speeding is typically faster than this prediction, leading researchers to ascribe the RSE to a so-called co-activation account. According to this account, multisensory neural processing occurs whereby the unisensory inputs are integrated to produce more effective sensory-motor activation. However, the typical paradigm used to test for RSE involves random sequencing of unisensory and bisensory inputs in a mixed design, raising the possibility of an alternate attention-switching account. This intermixed design requires participants to switch between sensory modalities on many task trials (e.g. from responding to a visual stimulus to an auditory stimulus). Here we show that much, if not all, of the RSE under this paradigm can be attributed to slowing of reaction times to unisensory stimuli resulting from modality switching, and is not in fact due to speeding of responses to AV stimuli. As such, the present data do not support a co-activation account, but rather suggest that switching and mixing costs akin to those observed during classic task-switching paradigms account for the observed RSE. © 2020 IBRO. Published by Elsevier Ltd. All rights reserved.

**Key words:** task-switching, attention, race model, redundant signals effect, multisensory, crossmodal.

## INTRODUCTION

A commonly deployed method for studying the behavioral effects of multisensory integration involves the use of a simple speeded reaction time (RT) task, whereby study participants are asked to respond as quickly as possible when presented with a unisensory stimulus (e.g. a tone or a briefly presented disk on a screen) or with combined bisensory stimuli (e.g. the synchronized presentation of a tone and a disk). In turn, the speed of responding to the bisensory (multisensory) input is compared to the response speeds to the constituent unisensory inputs. It has been consistently shown that

responses to such bisensory stimuli are substantially faster than responses to the unisensory inputs (Hershenson, 1962; Schroger and Widmann, 1998; Molholm et al., 2002; Megevand et al., 2013; Nidiffer et al., 2016). This has been termed the ‘redundant signals effect’ (RSE) and two plausible models have generally been forwarded to account for this speeding. The first is statistical facilitation (commonly referred to as the “race model” (Raab, 1962)). Under this account, each of the sensory signals in a bisensory pairing is independently processed and, on any particular trial, the RT is determined by whichever sensory system is first to trigger a response – the winner of the race, so to speak. The result of such a process over the course of multiple repetitions would lead to a bisensory RT distribution shifted toward shorter RTs relative to the distributions of either of the component unisensory RTs, but also one that is entirely predictable from said unisensory RT distributions. An alternate account, the so-called coactivation model (Miller, 1982), is based on the notion that the two initially

\*Correspondence to: J. J. Foxe, The Cognitive Neurophysiology Laboratory, The Del Monte Institute for Neuroscience, Department of Neuroscience, University of Rochester School of Medicine and Dentistry, Rochester, NY 14642, USA.

E-mail address: [john\\_foxe@urmc.rochester.edu](mailto:john_foxe@urmc.rochester.edu) (J. J. Foxe).

Abbreviations: AV, audio-visual; CDF, cumulative distribution function; RMV, race model violation; RSE, redundant signals effect; RT, reaction time.

separated sensory signals interact neurally prior to their detection, and that this integrated signal will tend to reach response threshold more quickly than either of the isolated signals. Under this account, multisensory speeding could potentially be faster than would be predicted from the unisensory RT distributions and would thereby “violate” the aforementioned race model.

The majority of prior RSE studies have supported the latter account, with few exceptions (Otto and Mamassian, 2012, 2017; Van der Stoep et al., 2015; Gondan and Minakata, 2016). To the extent of our knowledge, almost all multisensory tests of the race model in neurotypical adult populations have shown violation of the race model (Miller, 1982, 1986; Diederich and Colonius, 1987; Harrington and Peck, 1998; Molholm et al., 2002; Murray et al., 2005; Mahoney et al., 2015). As such, the field has generally embraced the notion that the RSE reflects multisensory neural integrative processes (i.e. the coactivation account) (Molholm et al., 2002; Foxe and Schroeder, 2005; Murray et al., 2005; Foxe and Molholm, 2009; Sperdin et al., 2009; Molholm and Foxe, 2010; Brandwein et al., 2011; Senkowski et al., 2011). This interpretation has been further bolstered by results from neurophysiological studies, which show evidence for multisensory integration during very early sensory processing (Molholm et al., 2002, 2006; Senkowski et al., 2006). That is, both electrophysiological and neuroimaging studies in humans and monkeys have suggested that audio-visual and audio-somatosensory inputs are integrated during the earliest phases of sensory processing, in hierarchically early cortical regions that were once thought to support only sensory-specific processing (Foxe et al., 2000, 2002; Schroeder and Foxe, 2002, 2005; Kayser et al., 2008; Boyle et al., 2017). Single unit studies in the superior colliculus of anesthetized cats have also shown enhanced responses to bisensory inputs, and of particular relevance here, these enhancements are greatest during the early phase of the response and show shortened neural response latencies (Rowland et al., 2007). While these results also suggest a potential coactivation account and a speeding of neural responses, it is important to point out that they have not been explicitly linked to behavioral outcomes.

A central design aspect, however, of almost all RSE studies in humans is that instances of unisensory and bisensory stimuli are intermixed randomly with variable inter-stimulus intervals (ISIs). In this way, participants cannot anticipate which sensory stimulus, or combination thereof, is about to occur and thus employ stimulus-specific processing strategies, and the use of highly variable ISIs counters the tendency for observers to try to anticipate the timing of the next stimulus. This random presentation pattern has been highlighted as a key aspect of such designs (Gondan and Minakata, 2016). However, there is an inherent issue with such mixing designs that calls into question a purely multisensory coactivation account for the RSE (Gondan et al., 2004). Every unisensory stimulus in a mixed block either triggers a repetition of a sensory-motor response mapping or a reconfiguration of the sensory-motor response mapping. That is, sometimes unisensory auditory inputs are fol-

lowed by another auditory input, and one could conceive of such instances as a “repeat” task. Sometimes, however, an auditory input is followed by a visual input (and vice versa), and here one might conceive of this as a “switch” task. Following this line of thinking, the traditionally used random presentation RSE design might be considered similar to more traditional task-switching paradigms (Rogers and Monsell, 1995).

If task switching, and thus switching attentional sets or task rules, is playing a prominent role in the emergence of the RSE under intermixed multisensory stimulation conditions, a number of simple predictions should hold. First, we should observe so-called ‘mixing costs’ (Rubin and Meiran, 2005; Wylie et al., 2009): the slowing of responses to each task in a task-switching design relative to “pure” blocks (i.e. blocks of stimuli where only one task is required). Second, we should observe ‘switching costs’ (Wylie et al., 2003b,a; 2004): the slowing of responses to, for example, auditory stimuli that immediately follow visual stimuli (and vice versa) compared to trial sequences in which stimuli are repeated. One would also predict that reaction times to the bisensory audio-visual stimulus should not be affected by these switching or mixing costs because they contain both stimulus types and therefore, no task ‘switch’ is strictly necessary.

Here, we test this alternate account of the RSE by testing the race model while accounting for the sequencing of stimuli (i.e. separating “switch” and “repeat” trials). We also conducted a set of “pure” block conditions (i.e. auditory-alone, visual-alone and audio-visual-alone stimulus sequences) to test for “mixing” costs by comparing RT patterns during these pure blocks to those observed during mixed randomly interleaved blocks of stimuli. If something akin to task-switching is present in classic RSE experimental designs, then it will be necessary to reassess the meaning and potential applications of race model analyses.

## EXPERIMENTAL PROCEDURES

### Subjects

Thirty-seven healthy adults (mean age = 28.3, SD = 6.6, 21 females, 16 males) participated. The Research Subjects Review Board of the University of Rochester approved the experimental procedures. Each participant provided written informed consent in accordance with the tenets laid out in the Declaration of Helsinki. Individuals were included in this study only if they were between the ages of 18 and 48 years, did not carry a diagnosis of a neurological or neurodevelopmental disorder, had healthy or corrected vision, were not hearing impaired, were not experiencing symptoms of post-concussive syndrome, were not dependent on drugs or alcohol, and were not under the influence of drugs or alcohol at the time of the study. Participants completed a questionnaire to confirm that inclusion criteria were met. Participants received a modest fee (\$14 per hour) for their efforts.

## Stimuli

The auditory (**A**) stimulus consisted of a 70 dB SPL, 1000 Hz tone (60 ms duration, 5 ms rise/fall times) presented over both channels of circumaural Sennheiser HD 600 headphones. The visual (**V**) stimulus consisted of a red disk presented for 60 ms, subtending 2.8 degrees in diameter at 85 cm viewing distance appearing on a dark gray background on an Acer Predator Z35 monitor. This stimulus was presented at an eccentricity of 2.9 degrees directly above a central fixation cross. The central fixation cross was present for the duration of all stimulus presentations. The audiovisual (**AV**) stimulus consisted of the simultaneous presentation of the above two stimuli.

## Procedure

Participants were seated comfortably in a darkened room and asked to respond as quickly as possible to each stimulus presentation. Responses to stimuli were made with a button press with the right thumb on an ergonomic Steelseries 3GC game controller. While this controller was not rigorously engineered for psychophysics experiments, in house testing demonstrated a precision of 9.7 ms between a button press and the registration of a button press by our stimulus delivery software.

Four block types were presented: unisensory-audio (pure **A**), unisensory-visual (pure **V**), bisensory-audiovisual (pure **AV**), and a mixed block of stimuli. The pure **A**, pure **V**, and pure **AV** blocks contained repetitions of only the **A**, **V**, or **AV** stimulus, respectively. Mixed blocks contained a randomly interleaved sequence of the three stimulus types (the classic RSE experimental design). This design and the parameters used are depicted in Fig. 1.

Mixed and pure blocks were randomly sequenced for each participant and 50 stimuli were presented in each block. Participants performed a minimum of 4 of each of the pure block types, for a total of 200 presentations of each of the pure stimuli. A minimum of 18 mixed blocks was presented, which resulted in approximately 100 presentations of each of the mixed block conditions. Throughout all blocks, ISIs were randomly varied between 1000 and 2400 ms with a square distribution to prevent anticipatory responding. Participants were encouraged to take as many breaks as desired between blocks to prevent fatigue. Stimuli were presented using Presentation software (Neurobehavioral Systems Inc., Berkeley, CA). Participants were asked to maintain fixation on the small, continuously presented fixation cross for all blocks, including pure **A** blocks. Furthermore, participants were explicitly instructed to press the button as quickly as possible for all stimulus conditions. Compliance with instructions was monitored during the experiment with a live stream from an EyeLink 1000 (SR Research, Ottawa, Canada) infrared eye-tracking camera. If participants were observed looking away from the screen, blinking excessively, or otherwise not participating according to instructions, they received further instruction from the investigators

and corresponding blocks were flagged for exclusion. Camera video data were not stored.

## Data acquisition and definitions

RT data were extracted from log files generated by the Presentation software and processed using custom, open-source MATLAB (MathWorks Inc., Natick, MA) scripts. Throughout this report we will refer to events presented in mixed blocks as either repeat or switch. That is, when the target stimulus is the same as the prior stimulus (i.e. **V** followed by **V**, **A** followed by **A**, and **AV** followed by **AV**), it is considered a repeat trial. When the target is different from the prior (i.e. **A** followed by **V**, **V** followed by **A**, and **A** or **V** followed by **AV**), it is considered a switch trial.

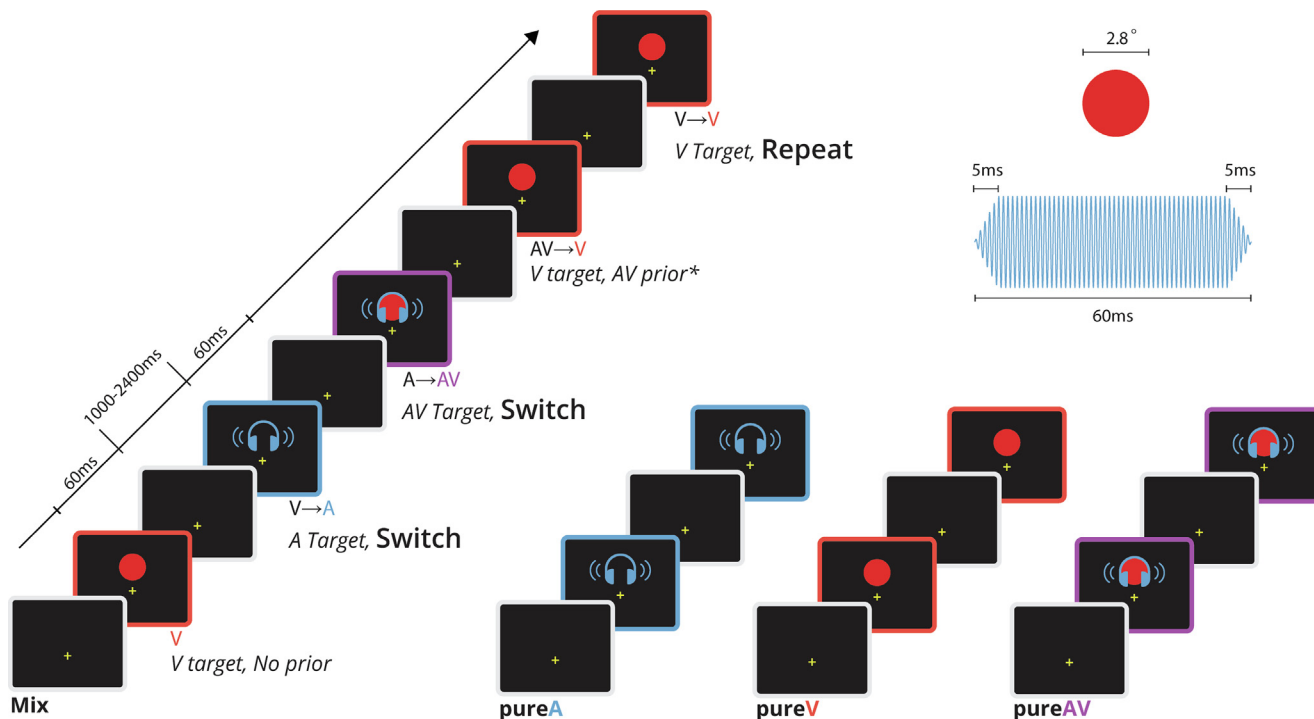
Stimuli from mixed blocks were also grouped without consideration of prior and target combinations. This is the “classic” mixed block approach to grouping stimuli from RSE paradigms, which we refer to as Mix **A**, Mix **V**, and Mix **AV**. We refer to stimuli grouped from pure blocks as Pure **A**, Pure **V**, and Pure **AV**. Pure, Mix, Repeat, and Switch comprise the main conditions of interest in our analyses. Mixed block **AV** stimuli were sorted into repeat and switch conditions for the purpose of comparison to unisensory stimuli. **AV** → **AV** was treated as a repeat. **A** → **AV** and **V** → **AV** are technically partial repeats and partial switches in a parallel processing account. For the sake of our analyses, we considered them to be switches and grouped them together to form a *switch AV* condition. **AV** → **A** and **AV** → **V** were grouped as neither a repeat nor a switch, as such a determination would, in an attentional shifting account, rely upon the stimulus presented prior to the **AV** stimulus. These latter conditions were included in the Mix **A** and Mix **V** groupings and included in pair-wise comparison tests described later.

## Data preparation

In order to reject outlier trials potentially due to undesirable response strategies, RTs outside of the middle 95th percentile of their respective conditions were excluded from analyses. This method avoided both arbitrary cutoff limits and nonparametric data cleaning approaches (Ulrich and Miller, 1994; Gondan and Minakata, 2016) while also accounting for both inter-participant and inter-condition variability. This data cleaning step was conducted prior to all analyses including distribution analyses.

## Reaction time and distribution analyses

In an initial survey of the RT data, distributions of the participant median RTs for each of the stimulus conditions were visualized with kernel density estimate plots, which can be seen in Fig. 2. Distributions of the median RTs for each condition were subjected to pair-wise, two-tailed nonparametric permutation tests with  $T_{max}$  multiple comparisons correction. This analysis was conducted in tandem with a distribution shift analysis



**Fig. 1.** Sensory modality and reaction time task. The task consisted of the sequential presentation of a red disc ( $2.8^\circ$  in diameter) on a monitor (V), a 1000 Hz tone played through headphones (A), or the simultaneous combination of the two (AV). Participants were instructed to respond as quickly as possible to any perceived stimulus with a button press. All stimulus durations were 60 ms. 1000–2400 ms elapsed between each stimulus. A fixation cross was persistently displayed on the monitor for all trials. Stimuli were presented in either mixed or pure blocks. Pure blocks contained repetitions of a single stimulus for the duration of the block. Mixed blocks contained a randomized sequence of **A**, **V**, and **AV** stimulus presentations. Mixed block conditions are defined by a combination of target stimulus and prior stimulus. These conditions can be grouped into repeats ( $A \rightarrow A$ ,  $V \rightarrow V$ ,  $AV \rightarrow AV$ ) and switches ( $V \rightarrow A$ ,  $A \rightarrow V$ ,  $A \rightarrow AV$ ,  $V \rightarrow AV$ ). \* $AV \rightarrow A$  and  $AV \rightarrow V$  do not fall exclusively within either of these groups.

(Rousselet et al., 2017), which compares the morphological features of two distributions. The shift test utilizes Harrell-Davis quantile estimations of two distributions as comparative anchor points to provide a description of the morphological differences between distributions (Wilcox, 1995; Wilcox et al., 2014). For each comparison of distributions, respective pairs of quantile values are compared using bootstrapped permutation tests and then subtracted to generate a shift function. 19 quantile points (5% to 95% in 5% increments) were compared for each distribution and comparisons utilized an alpha value of 0.05 with Hochberg multiple comparisons corrections (Hochberg, 1988).

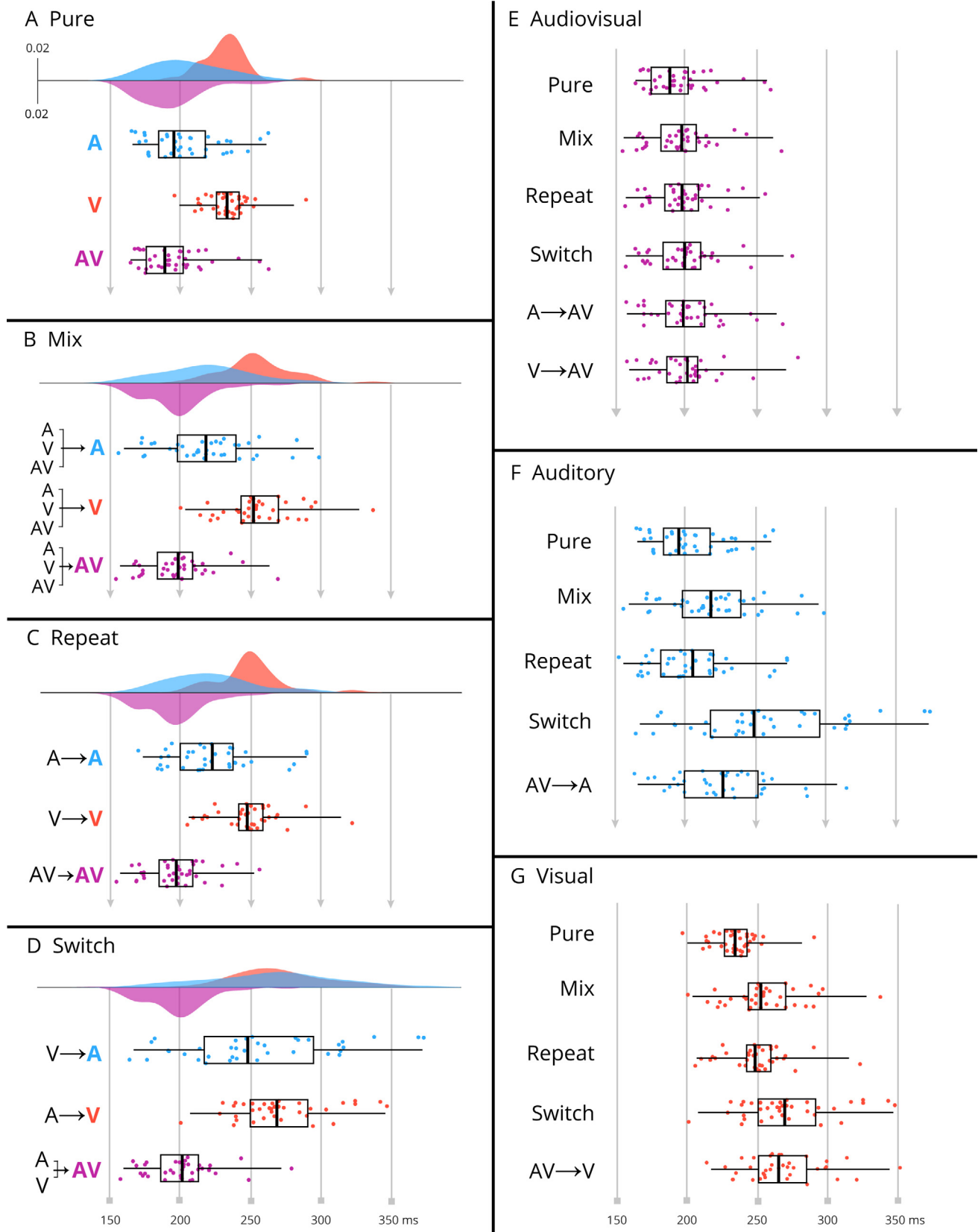
Mix cost values were generated for each participant for **A** and **V** conditions by subtracting median pure RT values from repeat RT values, respectively. Switch cost values were generated for **A** and **V** conditions for each participant by subtracting median repeat RT values from switch RT values for each sensory modality, respectively. Distributions of switch cost values were subjected to multiple comparison permutation tests using the abovementioned  $T_{max}$  correction method. All statistical analyses were performed in MATLAB: pairwise permutation tests were conducted using the PERMUTOOLS (Crosse et al., 2020a) open-source package available at <https://github.com/mickcrosse/PERMUTOOLS>, and distribution-shift tests (Rousselet et al.,

2017) were conducted using open-source code available at [https://github.com/GRousselet/matlab\\_stats](https://github.com/GRousselet/matlab_stats).

### Race model analyses

After data preparation, the maximum and minimum RT values for each participant across all conditions were used to quantize each participant's condition-wise RT distributions into 21 quantiles which were used to generate cumulative distribution functions (CDFs) and race models based on the unisensory CDFs. Condition-wise CDFs and race models were obtained by availing of all RTs within these upper and lower limits, except those that we considered outliers (i.e., RTs outside the middle 95th percentile). For visualization of RT distributions at the group level, CDFs and race models were averaged or Vincentized (Vincent, 1912; Ratcliff, 1979) over participants within each quantile and are illustrated here with bootstrapped 95% confidence intervals.

RTs were grouped based on their sequence (Pure, mix, repeat, switch) and used to generate and test separate race models within each group. The so-called "race model" is applied to test whether the observed RSE (in this case bisensory **AV** inputs) can be predicted by a simple probability summation account. If this is the case, then the following equation will hold true. Violations of the following equation imply that simple



**Fig. 2.** Participant median RT distributions. Distributions of participant median RT values are separated into race model analysis groups (**A** = Pure, **B** = Mix, **C** = Repeat, and **D** = Switch) and target stimuli (**E** = AudioVisual (magenta), **F** = Auditory (blue), and **G** = Visual (red)) to illustrate the effects of sequence conditions on RT distributions. Kernel density plots illustrate the density of population RT scatter plots. A and V are shown above the origin line. AV is inverted and shown below the origin line. All three plots depict positive density. Scatter plots show the distribution of participant median RTs. Each point depicts the median value of a participant for that condition. Box and whisker plots demonstrate the median values (bold central line) of each distribution and 2%, 25%, 75%, and 98% quantile points. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

parallel and independent unisensory stimulus processing is an insufficient account for reaction time phenomena in RSE paradigms.

$$P(\text{RT}_{\text{AV}} \leq t|\text{AV}) = P(\text{RT}_{\text{A}} \leq t|\text{A}) + P(\text{RT}_{\text{V}} \leq t|\text{V}) - P(\text{RT}_{\text{A}} \leq t|\text{A}) \times P(\text{RT}_{\text{V}} \leq t|\text{V})$$

We utilized the above equation to generate a race model prediction for each participant. This particular manifestation of the race model is known as Raab's model (Raab, 1962), which assumes independence between the **A** and **V** processing streams. Independence is statistically imparted by the subtraction of the joint probability term to the right of the equation. We used this model not necessarily because of its functional implications, but because the joint probability term reduces the built-in statistical facilitation of the race model prediction. This increases the likelihood of detecting multisensory speeding with respect to the model's prediction. Biasing the race model in this way presents a more conservative approach to testing for violation of the race model.

To validate the assumptions of our race model, a simulation of the race model was constructed for each participant by randomly selecting both **A** target RTs and **V** target RTs from their respective RT pools 1000 times. The minimum RT value was chosen for each pairing of **A** and **V** RTs and a CDF was generated from that distribution of minima. For each participant, this simulated CDF was compared to the race model function at each of the 21 quantile points using a permutation test with  $T_{\text{max}}$  correction. No significant differences were found at any quantile.

Violation of the race model was measured using two methods. The first method utilized  $T_{\text{max}}$  corrected permutation tests that compared the race function to the AV CDF at each quantile for each group. This yielded the identification of quantile ranges where the race model and AV distribution were statistically dissimilar. We utilize the  $T_{\text{max}}$  correction method (Westfall and Young, 1993; Blair et al., 1994) for all multiple comparisons analyses, as this has been shown to control for Type 1 error at a desired level and to be suitable in particular for testing the race model (Gondan, 2010). The second method entailed the estimation of race model violation (RMV), which was computed by subtracting the race model from AV CDF values at each quantile per group. The positive area contained within each participant's RMV curve was estimated using a trapezoidal integration function as it is common practice to interpret only the positive area under the curve (AUC) as an index of facilitative multisensory interactions (e.g., Miller, 1986; Nozawa et al., 1994). Participant RMV area measures were tested by group for deviation from zero with permutation tests. Note, inclusion of both the positive and negative AUC is qualitatively equivalent to using only the positive AUC (e.g., Nidiffer et al., 2016), because the positive and negative AUCs are inversely proportional (Crosse et al., 2019, Fig. 4B).

All race model analyses were conducted in MATLAB using routines based on the RaceModel open-source

toolbox (Crosse et al., 2020b), available at <https://github.com/mickcrosse/RaceModel>.

## RESULTS

### Median reaction time findings

Median RTs were determined for each participant for each experimental condition, and the resultant distributions were visualized using kernel density plots, which are presented alongside scatter plots with overlaid box and whisker plots in Fig. 2, panels A–D. RT statistics, including the median, median absolute deviation (MAD), which quantifies variance of the median, and the range of all conditions and groups, can be found in Table 1.  $T_{\text{max}}$  corrected p-values taken from the distribution comparison analysis and Cohen's *d* are reported in this section. Statistical comparisons between the conditions of main interest addressing this study's primary hypotheses are reported in what follows. However, a large number of between-condition comparisons was possible, and a comprehensive illustration and tabulation of all such combinations can be found in Fig. 3. While Fig. 3 shows the results of an additional non-parametric distribution shift test above the black unity line, we will report results only from the paired permutation tests. The purpose of the shift tests was to add an additional measure to interrogate the obvious morphological differences between the distributions seen in Fig. 2. The degree to which the results of these tests agree or disagree is discussed below.

In considering how RT is affected by stimulus (auditory, visual, audiovisual) and by condition (blocked, mix, repeat, and switch), we see in Table 1 that descriptively, **AV** RTs were uniformly the fastest, and that the fastest median RT here was in response to the AV stimulus in the Pure condition. Comparing the RTs across stimuli in the mix condition, **AV** RTs were statistically faster than **V** RTs ( $t = -27.23$ ,  $p = 0$ ,  $d = 2.33$ ) and **A** RTs ( $t = -8.39$ ,  $p = 0$ ,  $d = 0.73$ ). However, when pure condition stimuli were compared, **AV** RTs were found to be significantly faster than **V** RTs ( $t = -14.09$ ,  $p = 0$ ,  $d = 1.8$ ), whereas while there was a numeric difference for the **AV** versus **A** condition, this did not reach statistical significance ( $t = -2.18$ ,  $p = 0.16$ ,  $d = 0.23$ ). Fig. 2 illustrates the distributions of RTs across participants for each of the stimulus conditions.

Further parsing the RT data, Pure RTs were significantly faster than Repeat RTs (Pure **A** vs. Repeat **A**  $t = -4.52$ ,  $p = 0.0036$ ,  $d = -0.46$  | Pure **V** vs. Repeat **V**  $t = -8.89$ ,  $p = 0.00$ ,  $d = -0.93$ ) and Repeat RTs were faster than Mix RTs (Repeat **A** vs. Mix **A**  $t = 6.04$ ,  $p = 0.0001$ ,  $d = -0.29$  | Repeat **V** vs. Mix **V**  $t = 5.45$ ,  $p = 0.0003$ ,  $d = -0.48$ ). Furthermore, Mix RTs were faster than Switch RTs (Mix **A** vs. Switch **A**  $t = -5.47$ ,  $p = 0.0003$ ,  $d = -0.65$  | Mix **V** vs. Switch **V**  $t = -5.95$ ,  $p = 0.0002$ ,  $d = -0.43$ ). However, **AV** → **A** RTs were not statistically different from Repeat **A** RTs ( $t = -2.58$ ,  $p = 0.25$ ,  $d = 0.13$ ) and were statistically faster than Mix **A** RTs ( $t = 4.58$ ,  $p = 0.0031$ ,

**Table 1.** Distribution statistics

	Median	MAD	Min	Max	Avg Trials
Pure AV	186.6	17.0	159.0	257.4	204
Mix AV	199.3	16.5	155.0	270.2	227
Repeat AV	197.8	16.3	154.1	256.0	75
Switch AV	199.5	17.2	155.7	276.8	149
A → AV	197.0	18.5	153.5	269.0	74
V → AV	202.2	16.8	158.3	281.1	75
Pure A	192.5	20.7	158.1	255.3	208
Mix A	218.1	24.9	155.8	298.5	228
Repeat A	205.2	23.3	157.8	277.8	76
Switch A	239.1	42.2	156.4	364.7	76
AV → A	212.3	25.3	153.9	294.4	76
Pure V	227.4	11.3	188.9	282.6	205
Mix V	253.0	18.9	201.1	338.1	227
Repeat V	244.5	14.7	200.8	318.2	76
Switch V	263.8	25.1	199.0	344.9	74
AV → V	252.9	21.7	204.0	350.5	76

Participant population median RT distributions in milliseconds as described by median values, median absolute deviation (MAD), and the range of data (Min and Max). 37 participant medians for each of the above conditions were included in the generation of these population median values. The Average Trials column shows average trials used per participant after data cleaning.

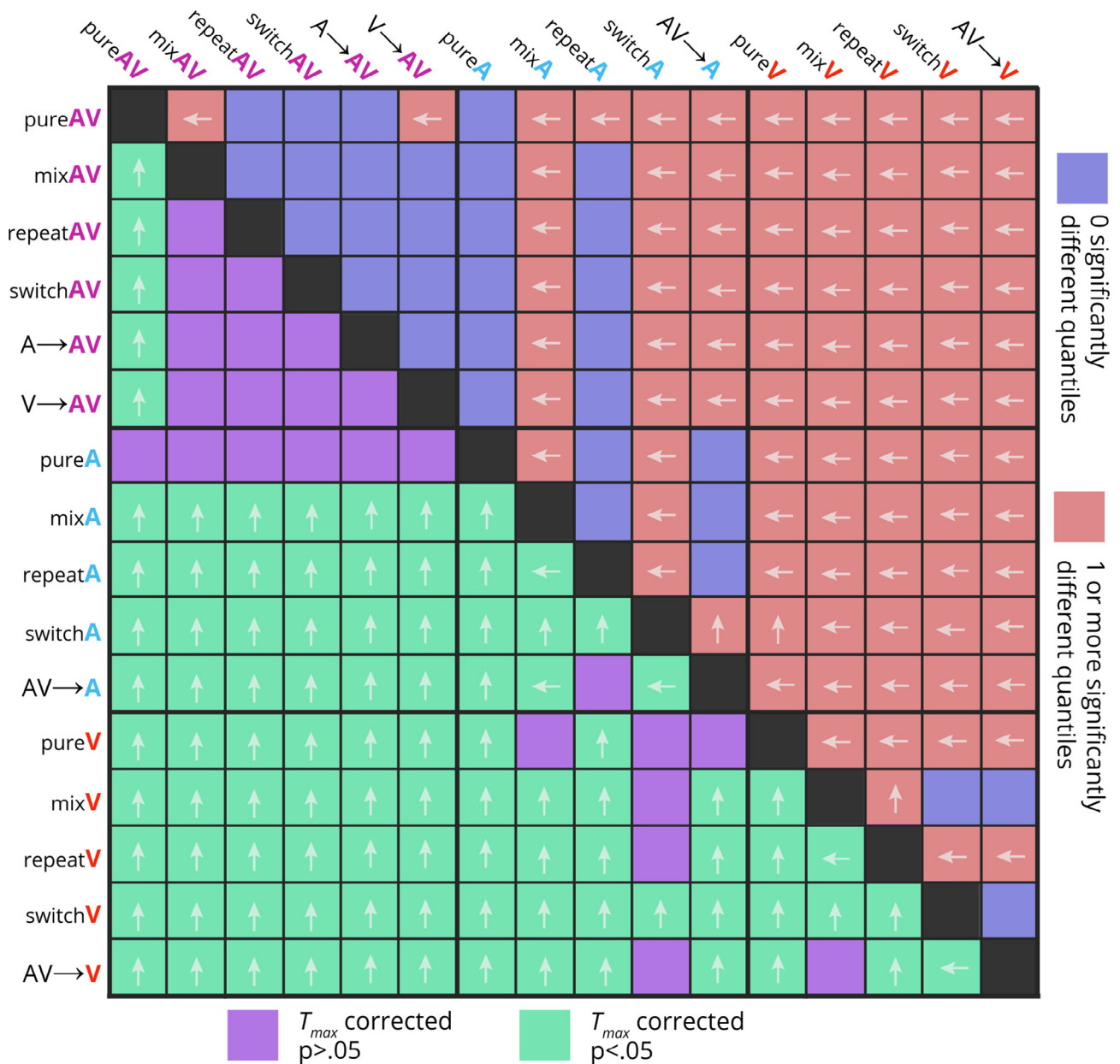
$d = -0.2$ ). On the other hand, **AV** → **V** RTs were slower than Repeat **V** ( $t = -4.68$ ,  $p = 0.0029$ ,  $d = 0.51$ ) and not statistically different from Mix **V** ( $t = -1.20$ ,  $p = 0.97$ ,  $d = 0.056$ ).

The lack of statistical difference in comparisons of mixed block **AV** conditions to other mixed block **AV** conditions makes it clear that responses to **AV** stimuli were less affected by the sequence of stimuli than were responses to **A** and **V** stimuli alone (Repeat **AV** vs. Mix **AV**  $t = 2.32$ ,  $p = 0.39$ ,  $d = -0.071$  | Mix **AV** vs. Switch **AV**  $t = -2.70$ ,  $p = 0.19$ ,  $d = -0.043$  | Repeat **AV** vs. Switch **AV**  $t = -2.56$ ,  $p = 0.26$ ,  $d = -0.11$ ). Nevertheless, the permutation tests show that Pure **AV** RTs were faster than Repeat **AV** RTs ( $t = -3.61$ ,  $p = 0.026$ ,  $d = -0.25$ ) and all other mixed-block-derived **AV** RTs (Pure **AV** vs. Mix **AV**  $t = -5.23$ ,  $p = 0.0005$ ,  $d = -0.31$  | Pure **AV** vs. Switch **AV**  $t = -5.82$ ,  $p = 0.0002$ ,  $d = -0.35$ ). However, it should be noted that the shift tests did not show significant differences in some of these comparisons of Pure **AV** to other **AV** conditions.

To summarize the incongruity between the two tests, the shift test found similarity while the permutation test found difference in comparisons of Pure **AV** to Repeat **AV**, Pure **AV** to Switch **AV**, and Pure **AV** to **A** → **AV**, comparisons of Repeat **A** to Mix **AV**, Repeat **A** to Repeat **AV**, Repeat **A** to Switch **AV**, Repeat **A** to **A** → **AV**, Repeat **A** to **V** → **AV**, and Repeat **A** to Mix **A**, comparisons of **AV** → **A** to Pure **A** and **AV** → **A** to Mix **A**, the comparison of Mix **V** to Switch **V**, and the comparison of Switch **V** to **AV** → **V**. The permutation test found similarity while the shift test found difference in comparisons of Pure **V** to Mix **A**, Pure **V** to **AV** → **A**, and Switch **A** to Pure **V**, Switch **A** to Mix **V**, Switch **A** to Repeat **V**, and Switch **A** to **AV** → **V**. Note that these latter comparisons are all cross-modality comparisons, which makes these results from the permutation test difficult to interpret. We found the distribution shift analysis to be an informative and robust measure of the

intrinsic characteristics of these RT distributions, but chose to base our interpretations on the more established comparison method. Mixing costs and switching costs are two prevalent features of task switching experiments, which we hypothesized would be observable in the RT distribution statistics and would subsequently affect race model results. We found a significantly greater magnitude of these costs in unisensory RTs relative to bisensory RTs. Compared to pure RTs, repeat **A** RTs were slowed by a  $7 \pm 11$  ms (median  $\pm$  MAD) mixing cost and **V** RTs were slowed by a  $14.8 \pm 10$  ms mixing cost. Additionally, compared to repeat RTs, switch **A** RTs were slowed by a  $27 \pm 31$  ms switching cost and **V** RTs were slowed by a  $24 \pm 20$  ms switching cost under respective switch conditions. **AV** RTs were affected to a lesser degree. **AV** RTs were slowed by a  $4 \pm 7$  ms mixing cost and by an additional  $3 \pm 5$  ms switching cost under switch conditions. A multiple comparisons permutation test was used to detect significant differences between stimuli for each cost type. Results in this section are  $T_{max}$  corrected  $p$ , Cohen's  $d$ , and each test has 36 degrees of freedom. The mixing cost test showed that the **AV** mixing costs were significantly less than both **A** mixing costs ( $t = -2.45$ ,  $p = 0.049$ ,  $d = 0.53$ ) and **V** mixing costs ( $t = -5.69$ ,  $p = 0$ ,  $d = -1.12$ ). **A** mixing costs were not significantly different than **V** mixing costs ( $t = -1.72$ ,  $p = 0.22$ ,  $d = -0.32$ ). Similarly, the **AV** switching costs were significantly less than both **A** switching costs ( $t = -5.45$ ,  $p = 0$ ,  $d = -1.24$ ) and **V** switching costs ( $t = -5.34$ ,  $p = 0$ ,  $d = -1.19$ ). **A** switching costs were significantly greater than **V** switching costs ( $t = 2.58$ ,  $p = 0.033$ ,  $d = 0.40$ ).

We would note that upon request of one of the reviewers of this manuscript, all above analyses were also performed on data where RTs were normalized at the individual participant level. This transformation had no impact on the results reported herein and is not detailed further. The authors are happy to provide these



**Fig. 3.** RT distribution multiple comparisons tests. Pairwise multiple comparison tests of participant median RT distributions were conducted using a permutation test with  $T_{max}$  multiple comparisons correction. Results for this test are shown below the black unity line. Comparisons yielding a corrected  $p < 0.05$  are colored green. Comparisons yielding a corrected  $p > 0.05$  are colored purple. Arrows indicate the group with the lowest median RT. The distribution shift test divided compared distributions into quantile-based comparison points and tested for difference at each quantile point. The results of this test are shown above the black unity line. Results for the shift test are quantified as the number of significantly different quantiles. Shift test results demonstrating one or more statistically different quantile point values qualify distributions as morphologically different and are colored red. Results demonstrating zero statistically different quantile point values qualify distributions as morphologically similar and are colored blue. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

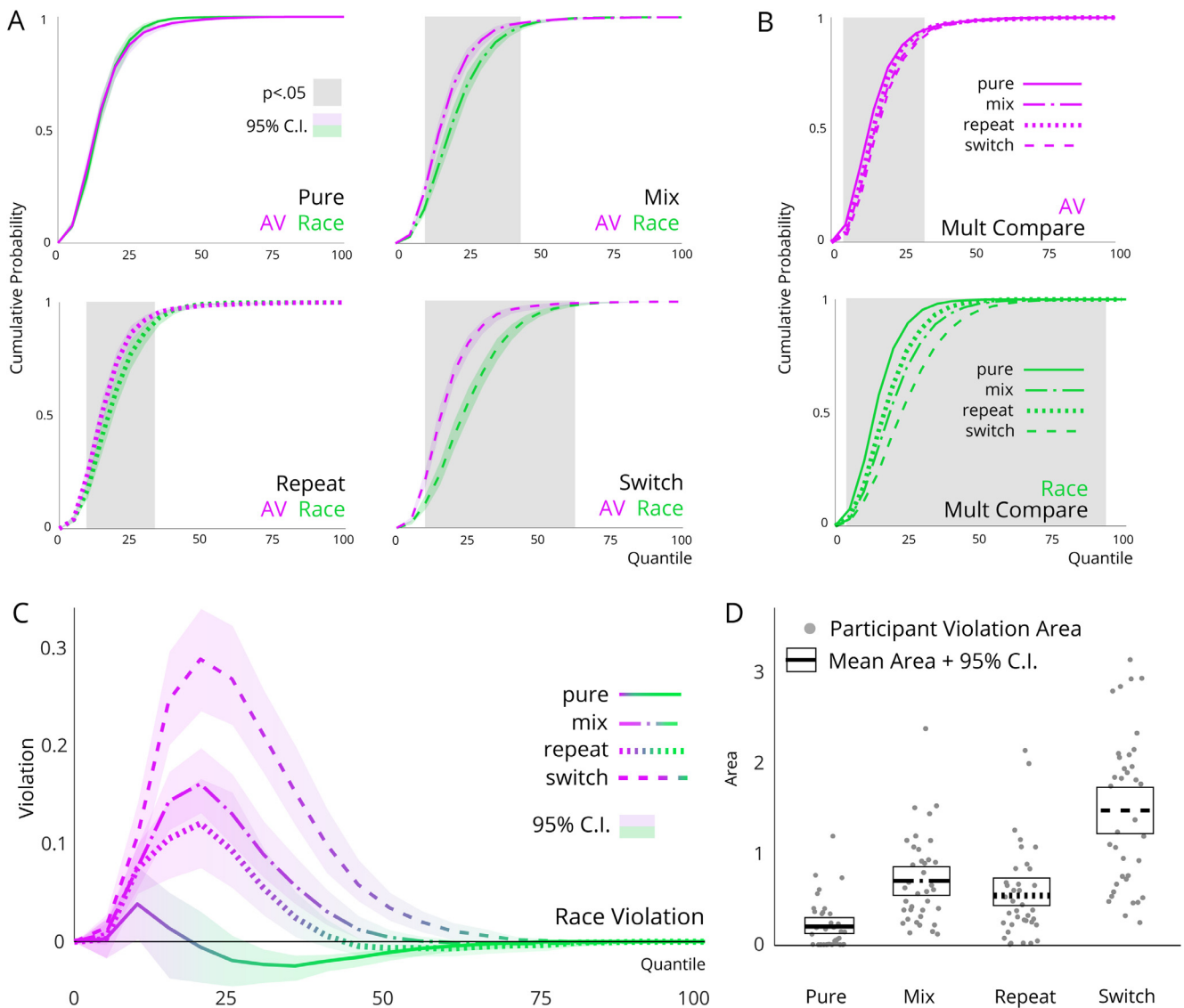
materials to any interested party under our data sharing agreement (see below).

**Race model findings**

Fig. 4A illustrates race model comparisons to respective AV CDFs for the pure, mix, repeat, and switch conditions over 21 quantiles. All permutation tests in this section were right tailed and conducted with 36 degrees of freedom, using  $T_{max}$  correction for multiple

comparisons.  $T$  values, corrected  $p$  values, original  $p$  values, and Cohen’s  $d$  values for each comparison can be found in Table 2. Despite consistent reports of race model violation in healthy adult populations, we found no statistical difference between the pure race model and the pure AV CDF at any quantile. The mix group comparison revealed seven significantly different quantiles at the 10%,15%, 20%, 25%, 30%, 35%, and 45% points (all  $p = 0.0$ ). The repeat group comparison revealed five significantly different quantiles (10–25%





**Fig. 4.** Race model quantile comparisons and inequality analysis. **(A)** Race model functions were compared to AV CDFs for each experimental group at 21 quantile points. Gray regions indicate significant quantiles of difference using  $T_{max}$  corrected  $p < 0.05$ . Colored regions indicate 95% confidence intervals (CI). **(B)** AV CDFs were subjected to multiple comparisons at each of the 21 quantile points. There were 6 possible combinations of group comparisons. Shaded regions indicate at least one significantly different comparison ( $T_{max}$  corrected  $p < 0.05$ ). This analysis was repeated for race model functions. **(C)** Race model violation (RMV) functions for each group were constructed by subtracting the AV CDF from the race model at each quantile point. These are shown with bootstrapped 95% CIs. **(D)** Scatter plots of the positive area contained beneath each participant's RMV function for each group with mean area  $\pm$  bootstrapped 95% CI. All group distributions were significantly different from one another and significantly non-zero ( $T_{max}$  corrected  $p < 0.05$ ).

$p = 0.0$ , 30%  $p = 0.01$ ). The switch group comparison revealed eleven significantly different quantiles (10–55%  $p = 0.0$ , 60%  $p = 0.04$ ). Thus, whereas the race model accurately approximated **AV** data from the pure condition, there was clear and progressively greater violation of the race model in the repeat, mix, and switch conditions.

Fig. 4B illustrates the comparison of the RT CDFs for AV stimuli across the different conditions and for the race model, which is composed of the unisensory RTs, across the different conditions. From this, it is evident that the AV distributions change very little while the race model CDF varies considerably as a function of context. Thus it is the unisensory distributions, which feed the race model,

that account for these contextual effects. We then asked whether the patterns of condition similarity and dissimilarity in the RT distribution analyses were apparent when race model functions were compared to themselves and **AV** CDFs were compared to themselves. Race model functions are composed from unisensory RT statistics; therefore we would expect to see differences between conditions in a similar fashion to the distribution comparisons. By extension, we would also expect to see less difference between the AV condition CDFs. To quantify this, we performed a quantile-wise multiple comparisons analysis on the race models for each condition and the **AV** CDFs for each condition, using two-tailed permutation tests with  $T_{max}$

**Table 2.** Comparisons of race models to AV CDFs

quantile	Pure Race v. Pure AV				Mix Race v. Mix AV				Repeat Race v. Repeat AV				Switch Race v Switch AV			
	<i>t</i>	Pcorr	Porig	<i>d</i>	<i>t</i>	Pcorr	Porig	<i>d</i>	<i>t</i>	Pcorr	Porig	<i>d</i>	<i>t</i>	Pcorr	Porig	<i>d</i>
0%	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
5%	0.28	0.50	0.39	0.05	1.36	0.41	0.09	0.26	0.29	0.50	0.39	0.06	2.20	0.08	0.02	0.46
10%	1.96	0.13	0.03	0.27	7.02	0.00	0.00	0.66	5.89	0.00	0.00	0.57	5.60	0.00	0.00	0.93
15%	0.60	0.50	0.28	0.08	9.30	0.00	0.00	0.81	5.14	0.00	0.00	0.51	9.93	0.00	0.00	1.46
20%	-0.31	0.50	0.62	-0.04	9.32	0.00	0.00	0.88	5.08	0.00	0.00	0.59	10.93	0.00	0.00	1.60
25%	-1.46	0.70	0.92	-0.18	7.31	0.00	0.00	0.86	3.99	0.00	0.00	0.59	9.82	0.00	0.00	1.57
30%	-2.71	0.98	1.00	-0.36	6.18	0.00	0.00	0.80	3.17	0.01	0.00	0.51	9.00	0.00	0.00	1.50
35%	-3.79	1.00	1.00	-0.60	5.00	0.00	0.00	0.74	2.17	0.11	0.02	0.38	7.41	0.00	0.00	1.37
40%	-4.44	1.00	1.00	-0.79	3.63	0.00	0.00	0.57	0.81	0.50	0.23	0.15	6.36	0.00	0.00	1.19
45%	-4.21	1.00	1.00	-0.83	2.21	0.09	0.01	0.34	-0.97	0.50	0.82	-0.15	5.52	0.00	0.00	1.00
50%	-3.84	1.00	1.00	-0.80	1.20	0.45	0.12	0.17	-2.07	0.87	0.98	-0.34	4.67	0.00	0.00	0.83
55%	-3.04	0.99	1.00	-0.65	0.20	0.50	0.42	0.03	-2.10	0.87	0.98	-0.37	3.77	0.00	0.00	0.70
60%	-2.68	0.97	1.00	-0.58	-0.38	0.50	0.64	-0.07	-2.31	0.92	0.99	-0.46	2.49	0.04	0.01	0.49
65%	-1.83	0.83	0.98	-0.41	-0.83	0.50	0.78	-0.17	-2.12	0.88	0.98	-0.46	1.90	0.18	0.03	0.40
70%	-1.52	0.73	0.94	-0.35	-1.42	0.61	0.91	-0.31	-1.87	0.80	0.94	-0.41	0.81	0.50	0.22	0.18
75%	-0.59	0.50	0.50	-0.14	-1.47	0.62	0.90	-0.33	-1.89	0.80	0.95	-0.42	0.67	0.50	0.26	0.15
80%	1.00	0.49	0.25	0.23	-1.58	0.66	0.93	-0.36	-1.29	0.56	0.75	-0.30	-0.45	0.50	0.63	-0.10
85%	NaN	NaN	NaN	NaN	-1.48	0.63	0.88	-0.33	-1.00	0.50	0.50	-0.23	0.05	0.50	0.49	0.01
90%	NaN	NaN	NaN	NaN	-0.78	0.50	0.50	-0.18	NaN	NaN	NaN	NaN	-0.25	0.50	0.50	-0.06
95%	1.43	0.31	0.00	0.00	1.76	0.23	0.00	0.41	NaN	NaN	NaN	NaN	1.74	0.23	0.00	0.40
100%	NaN	NaN	NaN	NaN	3.40	0.00	0.00	0.00	NaN	NaN	NaN	NaN	2.09	0.15	0.06	0.00

Race models compared to one another and AV CDFs compared to one another as seen in Fig. 3A. NaN indicates comparisons of distributions of equivalent values (i.e., entirely 0 at quantile 1 or entirely 1 in the final quantiles). Significant differences at the 100% quantile are ignored.

correction at each quantile. There were 6 possible inter-condition comparisons (pure vs. mix, pure vs. switch, pure vs. repeat, mix vs. switch, etc). Fig. 4B illustrates where at least one significantly different result was obtained with regions of shaded background. As can be seen, for both sets of comparisons, there are statistically significant differences as a function of condition. At the 5% quantile, there were 3 significantly different comparisons for the AV CDFs (pure vs. mix, pure vs. repeat, pure vs. switch) and also for the race model tests (pure vs. mix, pure vs. repeat, pure vs. switch). At 10%, all 6 inter-condition comparisons yielded significant differences for both the AV CDFs and the race model tests. At 15%, AV CDF comparisons yielded 5 significant differences (pure/mix, pure/repeat, pure/switch, mix/repeat, repeat/switch) and race model comparisons yielded significant differences for all 6 combinations. At 20–30%, AV CDF comparisons yielded 4 significant differences (pure vs. mix, pure vs. switch, mix vs. repeat, and repeat v. switch). All further comparisons of AV CDFs were not significantly different. On the other hand, all combinations of comparisons between the race model groups were significantly different from 20–50%. From 55–75%, all comparisons except for pure vs. repeat were significantly different for the race model comparisons. At 80%, race model pure vs. mix, pure vs. switch, mix vs. repeat, and rep vs. switch comparisons were significantly different. At 80%, all but pure v. repeat were significantly different. At 85%, race model pure vs. mix, pure vs. switch, mix vs. repeat, and repeat vs. switch were significantly different. At the 95% quantile point, no significant differences were found for all race model condition comparisons.

Finally, at the 100% quantile point, pure vs. mix and mix vs. repeat comparisons revealed significant differences. These differences had an effect size (Cohen's *d*) of 0 and therefore are not shown in the figure.

The condition-wise differences found in comparisons of race model and AV CDFs were further illustrated by race model violation (RMV) functions shown in Fig. 4C. The positive regions of these functions indicate regions where the AV CDF is "faster" than the race model prediction. The areas contained between these positive regions and the *x* axis were quantified for each condition and for each participant. The distribution of participant area values for each condition was tested for deviation from zero using a one-sample permutation test and compared to all other condition distributions with two-sample, two-tailed permutation tests with multiple comparisons  $T_{max}$  correction. All groups of area values were significantly different from one another (Pure vs. Mix  $t = 6.72$ ,  $p = 0$ ,  $d = 1.32$  | Pure vs. Repeat  $t = 4.10$ ,  $p = 0.0007$ ,  $d = 0.84$  | Pure vs. Switch  $t = 9.46$ ,  $p = 0$ ,  $d = 2.00$  | Mix vs. Repeat  $t = 2.81$ ,  $p = 0.025$ ,  $d = 0.33$  | Mix vs. Switch  $t = 8.53$ ,  $p = 0$ ,  $d = 1.10$  | Repeat vs. Switch  $t = 6.76$ ,  $p = 0$ ,  $d = 1.26$ ) and significantly non-zero (Pure  $t = 4.63$ ,  $p = 0$ ,  $d = 1.08$  | Mix  $t = 8.83$ ,  $p = 0$ ,  $d = 2.05$  | Repeat  $t = 6.67$ ,  $p = 0$ ,  $d = 1.55$  | Switch  $t = 10.57$ ,  $p = 0$ ,  $d = 2.46$ ). These results are illustrated in Fig. 4D. While it is the case that the pure AV CDF and pure race model were not significantly different from one another in the quantile comparison test (Fig. 3C, D), the pure condition violation area analysis does show a significantly non-zero distribution.

## DISCUSSION

In a mixed design, reaction times in response to bisensory stimuli are faster than responses to their constituent unisensory stimuli, giving rise to the redundant signals effect (RSE). One possible explanation of this finding is provided by the race model, which posits that bisensory RTs may be accounted for by simple statistically facilitation, i.e. the increased likelihood of having a fast RT. Under this account, no special multisensory interactive processing occurs; rather, on any given bisensory trial, the two inputs initiate independent sensory-motor processes, which “race” in parallel to initiate a response. However, ample evidence of speeding beyond the linear predictions of the race model has now been demonstrated, leading to a generally widespread acceptance of the co-activation model, which posits that multisensory integrative processes give rise to the observed RT speeding. The fact that RTs are often considerably faster than the predictions of the race model suggests that additional mechanisms must be at play. One plausible account for this is that the anatomical convergence of unisensory inputs on multisensory neurons leads to integration of sensory signals in a non-linear fashion, which in turn produce more effective sensory-motor processes, thereby resulting in additional speeding. Here, we build upon prior work (Gondan et al., 2004, 2007) that posits yet a third account for these multisensory RT phenomena, showing that much, if not all of the observed speeding can be explained in terms of classic mixing and switching costs that are more commonly associated with task-switching designs. That is, we find that the need to switch from performing this simple RT task in one sensory modality to performing the same task in a second sensory modality accounts for much of the bisensory “speeding” effects seen in these paradigms. The data here suggest that the main contributor to the observed RSE effects actually result from task-switching induced *slowing of unisensory responses*, rather than speeding of bisensory responses, which remain remarkably stable within the mixed block (i.e. **AV** RTs are wholly unaffected by whatever stimulus type preceded them).

It has long been appreciated that the order in which stimuli are presented in mixed blocks of stimuli has a tendency to slow RT behavior, and considerable empirical work has shown that RTs become systematically slower under both switching and mixing conditions (Rogers and Monsell, 1995; Spence et al., 2001; Wylie et al., 2003b; Gondan et al., 2004; Wylie et al., 2009; Sandhu and Dyson, 2013). The data here are strongly consistent with the presence of a mixing cost. That is, we observed robust slowing of RTs when stimuli were randomly presented in a mixed block compared to repeated presentation during pure blocks, even for responses to **AV** stimuli. Additionally, we showed that RTs are significantly slowed under switch conditions within the mixed block by comparing repeat and switch RTs. The presence of these two phenomena suggests that the brain treats the different stimulus modalities in RSE versus pure designs differently, with additional processing required to transition from one sensory modality

to another. Similar modality switching costs have been demonstrated and described previously in RSE designs (Gondan et al., 2004; Otto and Mamassian, 2012). Modality switching and mixing costs have also been demonstrated in cued and un-cued task-switching paradigms with sensory modality switching components (Rogers and Monsell, 1995; Sohn and Carlson, 2000; Wylie et al., 2009; Sandhu and Dyson, 2013).

It is worth re-emphasizing that bisensory RTs were not significantly affected by switch costs. As such, apparent multisensory speeding during mixed blocks appears to be fully generated by the slowing of responses to unisensory stimuli. When the causes of unisensory slowing (i.e., mixing and switching), were removed in the context of pure blocks, the race model provides a sufficient approximation of multisensory response times. It seems plausible that classical findings of multisensory speeding in RSE designs can be attributed, at least in part, to processes related to switching from one sensory modality to another (Spence et al., 2001; Shomstein and Yantis, 2004; Sandhu and Dyson, 2013). It is difficult to say what mechanisms generate these switching effects, but we are currently investigating the contributions of sensory attention, cross-sensory inhibition, and sensory motor mapping states.

Our choice to use pure blocks was motivated by the established use of repeated, “pure” task blocks in task switching designs. While the use of pure blocks is a common feature of task-switching designs, generating a race model from stimuli taken from separate types of blocks violates a well-established requirement for race model analyses known as context invariance (Miller, 1982; Ashby and Townsend, 1986; Luce, 1986; Gondan and Minakata, 2016). Race model analyses should assume that the environment and circumstances under which all data were collected were the same for all blocks and that participants were not utilizing alternate strategies that would invalidate race model comparisons. To put this in probabilistic terms, any probability distributions you wish to meaningfully compare must have the same priors. In the case of pure blocks, each separate block type (**A**, **V**, and **AV**) could be thought to generate different environmental circumstances that could motivate different response strategies, or isolate unique attentional circumstances or sets of sensory-motor mappings. To counteract the possibility of a participant adopting an overtly alternate strategy such as closing their eyes during an auditory block, for example, we monitored behavior in real time using an infrared camera system and provided corrective instruction when necessary. To our knowledge, the visual system was still monitoring the environment during the auditory pure blocks and vice versa. The overt difference between these blocks can thus likely be limited to attentional state. With this in mind, we note that it may be the case that responses to longer strings of repeats, i.e. **V** → **V** → **V**, in the mixed block will resemble pure RTs. Further experiments measuring mixing costs using multiple sequential repeat conditions will help to better understand these dynamics. Secondly, this method of grouping conditions in the mixed block (seen also in (Gondan et al., 2004)) may be a better application of the

race model test than former applications, which generally assumed that participants, while responding to stimuli in a mixed block, are in a uniformly multisensory attentional state.

While our experiments suggest that cross-sensory switching mechanisms account for the majority of race model violation that we see here in neurotypical adults (see Crosse et al., 2019 for an examination of the interplay of these factors in children and in individuals with autism), previous research shows that the relationship between unisensory slowing (switching) and multisensory speeding (coactivation) cannot be reduced to an either/or relationship, especially when stimuli become more complex or less detectable. For example, a similar pure block RSE analysis using auditory and visual stimuli (Otto and Mamassian, 2012) demonstrated a significant violation of the race model even when pure blocks of stimuli were used. The authors used stimuli that were embedded in noise, which may be a key difference between our experiments. The difference in our findings suggests that coactivation effects may become more apparent in more complex environments where unisensory stimuli have lower contrast from background. Although this experiment is part of a small pool of research to date (Otto and Mamassian, 2012) that takes switching and pure blocks into account using a race model analysis, there is a substantial body of inverse effectiveness research that supports the suggestion that stimuli with lower intensity or contrast can drive heightened multisensory neural responses (Stein et al., 1988; Ross et al., 2007; Senkowski et al., 2011; Crosse et al., 2016; Hong and Shim, 2016). In support of the present findings, a previous study that utilized a somewhat comparable pure block design to study multisensory processing during continuous (noise-free) audiovisual speech did not demonstrate race model violation when analyzing RTs to specific words (Crosse et al., 2015). However, in contrast to this body of research, recent work has shown evidence for coactivation in response to high intensity stimuli and not low intensity stimuli (Minakata and Gondan, 2018). It is likely the case that experimental context, stimulus choices, and subtle task choices have a great effect on the degree to which true multisensory speeding occurs. Deeper investigations of stimulus features such as intensity and contrast and their effects on cross-sensory processing have been conducted in murine models (Meijer et al., 2017) to suggest versatile feature-dependent activity modulation across separate primary sensory cortices. Studies of the temporal window of integration (Megevand et al., 2013) and even switch cost tuning as a function of inter-stimulus interval (Rogers and Monsell, 1995) suggest additional parameters that can be manipulated to better understand the relationship between putative coactivation effects and task-switching.

It is of significant interest to understand how an updated task-switching account of race model violation may contribute to developmental and clinical populations, from which a great deal of RSE data have been collected. Younger children do not show reliable multisensory speeding in classic RSE experiments, but evidence of speeding begins to emerge in later

childhood, at about the age of 10 (Brandwein et al., 2011, 2013; Crosse et al., 2019). This has been interpreted as evidence that multisensory integration processes have an extended developmental trajectory. Emergence of classic AV race model violation occurs even later in a number of clinical populations, such as those on the Autism Spectrum (Brandwein et al., 2013). Further work to translate these findings from the development of multisensory integration to multisensory attentional flexibility is in order (Reed and McCarthy, 2012; Crosse et al., 2019).

### Study limitations

One of the contentions in this paper is that no switching cost is observed in cases where a unisensory input is followed by the bisensory AV input, since the AV stimulus always contains the previously experienced input and so no switch of sensory modality is necessary. In the current study, the auditory and visual stimuli were of similar salience, and it is possible that this was a factor in the lack of switch costs we observed. That is, if one of the unisensory inputs were substantially more salient than the other, it remains possible that switch costs might emerge to AV inputs when the preceding input was the less salient one. This possibility will need to be tested in future work.

It has also been shown in prior work that the magnitude of the RSE can be affected by changes in the synchrony between the auditory and visual constituents of a bisensory input, especially when the salience of one of the unisensory inputs is substantially different to the other (Hershenson, 1962; Miller, 1986). We did not manipulate AV stimulus asynchrony here to maximize the RSE at the individual participant level, and it remains possible that evidence for co-activation might emerge under such circumstances. We would note, however, that in a prior study where we manipulated AV synchrony systematically using a near-identical stimulus setup to the one used in the current study (Megevand et al., 2013), the fastest responses to AV stimuli were observed when the audio and visual constituents of the bisensory input were presented simultaneously.

RSE designs have long been utilized to ask questions related to the nature and mechanisms of multisensory integration. Our analyses recast RSE designs into the realm of task-switching and provides evidence for putative attentional states that may drive the switching effects seen during alternation of responses to unisensory stimuli and the lack of switching effects found in response to bisensory stimuli. The analysis of pure block RTs demonstrates that, when taken out of the unpredictable context of a mixed block, responses to unisensory stimuli can be used to accurately approximate responses to bisensory stimuli. A great deal of work needs to be done to clarify the mechanisms that govern switching attention and spreading attention across modalities. Manipulations to stimulus intensity, contrast, complexity, predictability, and sensory modality could modulate cross-sensory attentional states and thus affect the resilience of multisensory stimuli to switch costs and to true MSI.

## ACKNOWLEDGMENTS

We thank our participants for their time and energy and the Cognitive Neurophysiology Laboratory (CNL) teams at Rochester and Einstein for their superb ongoing support of the work we do.

## FUNDING

This work was supported through funds provided to JJF by the University of Rochester Medical Center. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## DATA SHARING

Upon acceptance for publication, the authors will coordinate with the editorial office to ensure that the full datasets are uploaded to a publicly available repository and that the appropriate link is made to the article file. As mentioned in the main text, we will be happy to provide additional analyses that were conducted (such as those using normalized data) upon request.

## AUTHOR CONTRIBUTIONS

JJF, SM, EGF and LHS designed the experiment. Data were collected by LHS, EN, AMC and MSB. Analyses and illustrations were produced by LHS, EN and MJC. LHS produced the first draft of the manuscript. Multiple rounds of editing occurred with major contributions to this process from MJC, SM, EGF and JJF. All authors read and approved the final version of this manuscript.

## COMPLIANCE WITH ETHICAL STANDARDS

The Research Subjects Review Board of the University of Rochester approved all the experimental procedures. Each participant provided written informed consent in accordance with the tenets laid out in the Declaration of Helsinki.

## REFERENCES

- Ashby FG, Townsend JT (1986) Varieties of perceptual independence. *Psychol Rev* 93:154.
- Blair RC, Higgins JJ, Karniski W, Kromrey JD (1994) A study of multivariate permutation tests which may replace Hotelling's T2 test in prescribed circumstances. *Multivariate Behavioral Res* 29:141–163.
- Boyle SC, Kayser SJ, Kayser C (2017) Neural correlates of multisensory reliability and perceptual weights emerge at early latencies during audio-visual integration. *Eur J Neurosci* 46:2565–2577.
- Brandwein AB, Foxe JJ, Butler JS, Russo NN, Altschuler TS, Gomes H, Molholm S (2013) The development of multisensory integration in high-functioning autism: high-density electrical mapping and psychophysical measures reveal impairments in the processing of audiovisual inputs. *Cereb Cortex* 23:1329–1341.
- Brandwein AB, Foxe JJ, Russo NN, Altschuler TS, Gomes H, Molholm S (2011) The development of audiovisual multisensory integration across childhood and early adolescence: a high-density electrical mapping study. *Cereb Cortex* 21:1042–1055.
- Crosse MJ, Butler JS, Lalor EC (2015) Congruent visual speech enhances cortical entrainment to continuous auditory speech in noise-free conditions. *J Neurosci* 35:14195–14204.
- Crosse MJ, Di Liberto GM, Lalor EC (2016) Eye can hear clearly now: inverse effectiveness in natural audiovisual speech processing relies on long-term crossmodal temporal integration. *J Neurosci* 36:9888–9895.
- Crosse MJ, Foxe JJ, Molholm S. (2019) Developmental Recovery of Impaired Multisensory Processing in Autism and the Cost of Switching Sensory Modality. 565333.
- Crosse MJ, Foxe JJ, Molholm S. (2020a) PERMUTOOLS: A MATLAB Package for Multivariate Permutation Testing.
- Crosse MJ, Foxe JJ, Molholm S. (2020b) RaceModel: A MATLAB Package for Stochastic Modelling of Multisensory Reaction Times.
- Diederich A, Colonius H (1987) Intersensory facilitation in the motor component? *Psychol Res* 49:23–29.
- Foxe JJ, Molholm S (2009) Ten years at the multisensory forum: musings on the evolution of a field. *Brain Topography* 21:149–154.
- Foxe JJ, Morocz IA, Murray MM, Higgins BA, Javitt DC, Schroeder CE (2000) Multisensory auditory-somatosensory interactions in early cortical processing revealed by high-density electrical mapping. *Brain Res Cogn Brain Res* 10:77–83.
- Foxe JJ, Schroeder CE (2005) The case for feedforward multisensory convergence during early cortical processing. *Neuroreport* 16:419–423.
- Foxe JJ, Wylie GR, Martinez A, Schroeder CE, Javitt DC, Guilfoyle D, Ritter W, Murray MM (2002) Auditory-somatosensory multisensory processing in auditory association cortex: an fMRI study. *J Neurophysiol* 88:540–543.
- Gondan M (2010) A permutation test for the race model inequality. *Behav Res Methods* 42:23–28.
- Gondan M, Lange K, Rosler F, Roder B (2004) The redundant target effect is affected by modality switch costs. *Psychon B Rev* 11:307–313.
- Gondan M, Minakata K (2016) A tutorial on testing the race model inequality. *Atten Percept Psychophys* 78:723–735.
- Gondan M, Vorberg D, Greenlee MW (2007) Modality shift effects mimic multisensory interactions: an event-related potential study. *Exp Brain Res* 182:199–214.
- Harrington LK, Peck CK (1998) Spatial disparity affects visual-auditory interactions in human sensorimotor processing. *Exp Brain Res* 122:247–252.
- Hershenson M (1962) Reaction time as a measure of intersensory facilitation. *J Exp Psychol* 63:289–293.
- Hochberg Y (1988) A sharper Bonferroni procedure for multiple tests of significance. *Biometrika* 75:2.
- Hong SW, Shim WM (2016) When audiovisual correspondence disturbs visual processing. *J Exp Brain Res* 234:1325–1332.
- Kayser C, Petkov CI, Logothetis NK (2008) Visual modulation of neurons in auditory cortex. *Cereb Cortex* 18:1560–1574.
- Luce RD (1986) Response times: their role in inferring elementary mental organization. Oxford University Press on Demand.
- Mahoney JR, Molholm S, Butler JS, Sehatpour P, Gomez-Ramirez M, Ritter W, Foxe JJ (2015) Keeping in touch with the visual system: spatial alignment and multisensory integration of visual-somatosensory inputs. *Front Psychol* 6:1068.
- 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 e71608.
- Meijer GT, Montijn JS, Pennartz CMA, Lansink CS (2017) Audiovisual modulation in mouse primary visual cortex depends on cross-modal stimulus configuration and congruency. *J Neurosci* 37:8783–8796.
- Miller J (1982) Divided attention: evidence for coactivation with redundant signals. *Cogn Psychol* 14:247–279.
- Miller J (1986) Timecourse of coactivation in bimodal divided attention. *Percept Psychophys* 40:331–343.

- Minakata K, Gondan M (2018) Differential coactivation in a redundant signals task with weak and strong go/no-go stimuli. *Q J Exp Psychol (Hove)*. 1747021818772033.
- Molholm S, Foxe JJ (2010) Making sense of multisensory integration. *Eur J Neurosci* 31:1709–1712.
- 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* 14:115–128.
- Molholm S, Sehatpour P, Mehta AD, Shpaner M, Gomez-Ramirez M, Ortigue S, Dyke JP, Schwartz TH, Foxe JJ (2006) Audio-visual multisensory integration in superior parietal lobule revealed by human intracranial recordings. *J Neurophysiol* 96:721–729.
- Murray MM, Molholm S, Michel CM, Heslenfeld DJ, Ritter W, Javitt DC, Schroeder CE, Foxe JJ (2005) Grabbing your ear: rapid auditory-somatosensory multisensory interactions in low-level sensory cortices are not constrained by stimulus alignment. *Cereb Cortex* 15:963–974.
- Nidiffer AR, Stevenson RA, Krueger Fister J, Barnett ZP, Wallace MT (2016) Interactions between space and effectiveness in human multisensory performance. *Neuropsychologia* 88:83–91.
- Nozawa G, Reuter-Lorenz PA, Hughes HC. Parallel and serial processes in the human oculomotor system: bimodal integration and express saccades. *Biol Cybern*. 1994;72(1):19–34.
- Otto TU, Mamassian P (2012) Noise and correlations in parallel perceptual decision making. *Curr Biol* 22:1391–1396.
- Otto TU, Mamassian P (2017) Multisensory decisions: the test of a race model, its logic, and power. *Multisensory Res*:30.
- Raab DH (1962) Statistical facilitation of simple reaction times. *Trans N Y Acad Sci* 24:574–590.
- Ratcliff R (1979) Group reaction time distributions and an analysis of distribution statistics. *Psychol Bull* 86:446–461.
- Reed P, McCarthy J (2012) Cross-modal attention-switching is impaired in autism spectrum disorders. *J Autism Dev Disorders* 42:947–953.
- Rogers RD, Monsell S (1995) Costs of a predictable switch between simple cognitive tasks. *J Exp Psychol: General* 124:24.
- Ross LA, Saint-Amour D, Leavitt VM, Javitt DC, Foxe JJ (2007) Do you see what I am saying? Exploring visual enhancement of speech comprehension in noisy environments. *Cereb Cortex* 17:1147–1153.
- Rousselet GA, Pernet CR, Wilcox RR. Beyond differences in means: robust graphical methods to compare two groups in neuroscience. *Eur J Neurosci*. 2017 Jul;46(2):1738–1748. doi: 10.1111/ejn.13610.
- Rowland BA, Quessy S, Stanford TR, Stein BE (2007) Multisensory integration shortens physiological response latencies. *J Neurosci* 27:5879–5884.
- Rubin O, Meiran N (2005) On the origins of the task mixing cost in the cuing task-switching paradigm. *J Exp Psychol Learn Mem Cogn* 31:1477–1491.
- Sandhu R, Dyson BJ (2013) Modality and task switching interactions using bi-modal and bivalent stimuli. *Brain Cogn* 82:90–99.
- Schroeder CE, Foxe J (2005) Multisensory contributions to low-level, 'unisensory' processing. *Curr Opin Neurobiol* 15:454–458.
- Schroeder CE, Foxe JJ (2002) The timing and laminar profile of converging inputs to multisensory areas of the macaque neocortex. *Brain Res Cogn Brain Res* 14:187–198.
- Schroger E, Widmann A (1998) Speeded responses to audiovisual signal changes result from bimodal integration. *Psychophysiology* 35:755–759.
- Senkowski D, Molholm S, Gomez-Ramirez M, Foxe JJ (2006) Oscillatory beta activity predicts response speed during a multisensory audiovisual reaction time task: a high-density electrical mapping study. *Cereb Cortex* 16:1556–1565.
- Senkowski D, Saint-Amour D, Hofle M, Foxe JJ (2011) Multisensory interactions in early evoked brain activity follow the principle of inverse effectiveness. *Neuroimage* 56:2200–2208.
- Shomstein S, Yantis S (2004) Control of attention shifts between vision and audition in human cortex. *J Neurosci* 24:10702–10706.
- Sohn MH, Carlson RA (2000) Effects of repetition and foreknowledge in task-set reconfiguration. *J Exp Psychol Learn Mem Cogn* 26:1445–1460.
- Spence C, Nicholls ME, Driver J (2001) The cost of expecting events in the wrong sensory modality. *Percept Psychophys* 63:330–336.
- Sperdin HF, Cappe C, Foxe JJ, Murray MM (2009) Early, low-level auditory-somatosensory multisensory interactions impact reaction time speed. *Front Integr Neurosci* 3:2.
- Stein BE, Huneycutt WS, Meredith MA (1988) Neurons and behavior: the same rules of multisensory integration apply. *Brain Res* 448:355–358.
- Ulrich R, Miller J (1994) Effects of truncation on reaction time analysis. *J Exp Psychol Gen* 123:34–80.
- Van der Stoep N, Spence C, Nijboer TC, Van der Stigchel S (2015) On the relative contributions of multisensory integration and crossmodal exogenous spatial attention to multisensory response enhancement. *Acta Psychol (Amst)* 162:20–28.
- Vincent SB (1912) The function of the vibrissae in the behavior of the white rat. *Behavior Monographs* 1.
- Westfall PH, Young SS. (1993) *Resampling-Based Multiple Testing: Examples and Methods for p-Value Adjustment (Wiley Series in Probability and Statistics)*.
- Wilcox RR (1995) Comparing two independent groups via multiple quantiles. *J Royal Stat Soc, Series D (The Statistician)* 8.
- Wilcox RR, Erceg-Hurn DM, Clark F, Carlson M (2014) Comparing two independent groups via the lower and upper quantiles. *J Stat Comput Sim* 84:8.
- Wylie GR, Javitt DC, Foxe JJ (2003a) Cognitive control processes during an anticipated switch of task. *Eur J Neurosci* 17:667–672.
- Wylie GR, Javitt DC, Foxe JJ (2003b) Task switching: a high-density electrical mapping study. *Neuroimage* 20:2322–2342.
- Wylie GR, Javitt DC, Foxe JJ (2004) The role of response requirements in task switching: dissolving the residue. *Neuroreport* 15:1079–1087.
- Wylie GR, Murray MM, Javitt DC, Foxe JJ (2009) Distinct neurophysiological mechanisms mediate mixing costs and switch costs. *J Cogn Neurosci* 21:105–118.

(Received 24 October 2019, Accepted 6 April 2020)  
(Available online 20 April 2020)