



# Self-prioritization with unisensory and multisensory stimuli in a matching task

Clea Desebrock<sup>1</sup> · Charles Spence<sup>1</sup> · Ayla Barutçu<sup>1</sup>

Accepted: 21 April 2022 / Published online: 10 May 2022  
© The Author(s) 2022

## Abstract

A shape-label matching task is commonly used to examine the self-advantage in motor reaction-time responses (the Self-Prioritization Effect; SPE). In the present study, auditory labels were introduced, and, for the first time, responses to unisensory auditory, unisensory visual, and multisensory object-label stimuli were compared across block-type (i.e., trials blocked by sensory modality type, and intermixed trials of unisensory and multisensory stimuli). Auditory stimulus intensity was presented at either 50 dB (Group 1) or 70 dB (Group 2). The participants in Group 2 also completed a multisensory detection task, making simple speeded motor responses to the shape and sound stimuli and their multisensory combinations. In the matching task, the SPE was diminished in intermixed trials, and in responses to the unisensory auditory stimuli as compared with the multisensory (visual shape+auditory label) stimuli. In contrast, the SPE did not differ in responses to the unisensory visual and multisensory (auditory object+visual label) stimuli. The matching task was associated with multisensory ‘costs’ rather than gains, but response times to self- versus stranger-associated stimuli were differentially affected by the type of multisensory stimulus (auditory object+visual label or visual shape+auditory label). The SPE was thus modulated both by block-type and the combination of object and label stimulus modalities. There was no SPE in the detection task. Taken together, these findings suggest that the SPE with unisensory and multisensory stimuli is modulated by both stimulus- and task-related parameters within the matching task. The SPE does not transfer to a significant motor speed gain when the self-associations are not task-relevant.

**Keywords** Self-prioritization · Matching · Simple detection · Multisensory · Self-relevance · Auditory labels · Blocked vs. intermixed

## Introduction

Self-representation is widely held to guide our cognition and action, and self-relevance has repeatedly been shown to

influence stimulus processing (Cunningham & Turk, 2017; Sui & Humphreys, 2017b). Using stimuli such as our own face or name, our attributes (e.g., personality traits), our property, and also those stimuli that are newly and arbitrarily associated with the self, studies have investigated self-relevance across a diverse range of tasks. Self-related stimuli have been shown to modulate attention (Alexopoulos et al., 2012; Brédart et al., 2006; Humphreys & Sui, 2016; Moray, 1959), perceptual decision-making (Constable et al., 2019; Golubickis et al., 2017; Humphreys & Sui, 2016), memory (Rogers et al., 1977; Symons & Johnson, 1997; Yin et al., 2019), and overt movements (Constable et al., 2011; Constable et al., 2014; Desebrock et al., 2018; Desebrock & Spence, 2021).

A paradigm for investigating the effects of self-relevance without the confounds of stimulus familiarity (inherent in own-name or -face stimuli, for example) was established with the introduction of Sui et al.’s (2012) matching procedure. In an initial learning phase, participants are instructed to associate geometric shapes with visually presented labels referring

## Key Findings

- Label and object modality (auditory and visual) modulated the Self-Prioritization Effect (SPE) in a multisensory adaptation of a standard matching task.
- The SPE was diminished in intermixed as compared with blocked modality trials.
- Auditory stimulus intensity (50 dB vs. 70 dB) had no effect on the SPE.
- There was no SPE in a multisensory simple detection motor response paradigm.
- The SPE can be modulated by both stimulus- and task-related parameters in the matching task.

✉ Clea Desebrock  
Clea.Desebrock@psy.ox.ac.uk

<sup>1</sup> Department of Experimental Psychology, Crossmodal Research Laboratory, University of Oxford, Oxford OX2 6GG, UK

to people (e.g., self-circle, friend-triangle, stranger-square). Arbitrary associations between the identities and the stimuli are thus formed rapidly. In the main task, participants then indicate whether shape-label stimulus pairs match or mismatch the newly learned associations, typically by means of keypress responses. Reaction times (RTs) to the self-associated shape-label stimulus pairs are consistently found to be shorter and more accurately selected. This phenomenon is known as the Self-Prioritization Effect (SPE; Sui et al., 2012). Self-prioritization can thus be measured without the need to use stimulus objects that are highly familiar. Self-prioritization appears to be dissociable (at least in part) from the effects of stimulus familiarity, emotional valence, and reward (Schäfer et al., 2020a; Stolte et al., 2017; Stolte et al., 2021; Sui et al., 2012; Woźniak & Knoblich, 2019; Yankouskaya et al., 2017). Furthermore, distinct neural circuitry underpins the SPE (Sui et al., 2013; Yankouskaya et al., 2017).

### The effect of sensory modality on the Self-Prioritization Effect (SPE)

Typically, the object and label stimuli used in the matching task have been visual (e.g., Golubickis et al., 2017; Sui et al., 2012; Woźniak et al., 2018), and the SPE has been interpreted as arising from the binding of the visual self-associated stimulus to the self-concept (Humphreys & Sui, 2015; Schäfer et al., 2020a). However, it has previously been assumed (Humphreys & Sui, 2016) and later evidenced, in studies using auditory and tactile objects, that the SPE in the matching task is not visual-specific (Payne et al., 2020; Schäfer et al., 2015; Schäfer et al., 2016b; Stolte et al., 2021). Furthermore, it has been suggested that the self-representation underpinning self-prioritization may be modality-general (or ‘abstract’; Woźniak et al., 2018). Indeed, Schäfer et al. (2016b) documented that the magnitude of the SPE did not appear to differ in responses to visual as compared with auditory and tactile object stimuli (with objects presented 500 ms before the visual label stimuli), suggesting that the SPE was underpinned by neural processes that are common to responses across all sensory stimuli. The authors concluded that the SPE was therefore likely to be a modality-general mechanism.

In contrast, Stolte et al. (2021) used simultaneous presentations of auditory objects and visual labels (as such creating the opportunity for multisensory integration) and documented that the SPE was smaller than to the standard visual shape and visual label stimulus pairs. The authors suggested that under some conditions vision may dominate for self-associations (Stolte et al., 2021). They also found that the SPE and visual dominance can be modulated by the relative frequencies of the auditory tones used to pair with the self-associated labels. As

Hutmacher (2019) points out, visual dominance is socially and culturally reinforced. We tend to rely more on visual than auditory (or other sensory) information (Sinnott et al., 2007), and actively attend to visual over auditory stimuli in certain task contexts (Posner et al., 1976). Any task-induced visual dominance may be expected to reduce the SPE with auditory stimuli.

### Task design, stimulus parameters, and the SPE

Perhaps unsurprisingly, certain stimulus parameters and elements of the task design differed across the previous studies that have compared the SPE in responses to both visual and auditory stimuli (Schäfer et al., 2016b; Stolte et al., 2021). For example, Schäfer et al. presented their auditory stimuli at 50 dB, and used a blocked-trial design (with responses to auditory and visual stimuli being assessed in separate experiments). These authors documented that the magnitude of the SPE did not differ in responses to visual as compared to auditory stimuli. By contrast, Stolte et al. presented their auditory stimuli at 75 dB, and used an intermixed presentation with all stimulus modality types randomly intermixed within a block of experimental trials. These authors documented that the SPE in responses to paired auditory object and visual label stimuli was smaller than to the standard visual shape and visual label stimulus pairs. Such differences in the stimulus parameters and task design across studies may differentially influence the SPE across visual and auditory stimuli.

**Blocked versus intermixed trials** Task design (i.e., whether trials are intermixed or blocked by the identity of the target shape) has previously been demonstrated to influence self-prioritization. In a recent study using visual stimuli in the matching task, the authors documented that the SPE was greater in intermixed trials (using self- and friend-associated shapes) than in blocked trials (i.e., using self- or friend-associated shapes; Golubickis & Macrae, 2021). Blocking or intermixing trials by stimulus modality may also modulate the SPE. In intermixed as compared with blocked trials, responding to the randomised unisensory and multisensory trials requires modality-switching. Modality-switching typically results in slower and less accurate responses (Spence et al., 2001), reflecting ‘switching costs’ (Barutchu & Spence, 2021; Otto & Mamassian, 2012). Sensory switch costs often result in slower responses to unisensory stimuli since attention must be shifted between the senses (Kreutzfeldt et al., 2015; Lukas et al., 2010; Spence et al., 2001). In contrast, responses to multisensory stimuli require no switching since processing just one of the two sensory stimuli is sufficient to produce a valid motor response in those detection paradigms with redundant signals. Switching can

thus inflate multisensory gains in detection paradigms (Barutchu & Spence, 2021; Otto & Mamassian, 2012; Shaw et al., 2020). In a matching task, on the other hand, both signals need to be processed in order to make a decision and thereafter a valid motor response; therefore a switch to a multisensory signal in a matching task may be expected to lead to a multisensory cost. Indeed, multisensory costs were observed in Stolte et al.'s (2021) study, along with a reduced SPE in responses to those audiovisual stimuli. Switching could thus (partly) underpin the reduced SPE with audiovisual stimuli in their study, which was further investigated in the present study.

Increased demands on working memory in intermixed trials, as compared with blocked trials, may also favour those responses involving the stronger of the object-label associations (Golubickis & Macrae, 2021); in other words, those involving visual self-associations (should they be stronger than their auditory counterparts). Certain self-related categories appear to be prioritized more than others when placed into competition (Enock et al., 2020; Turner et al., 1987); for example, in intermixed trials (Enock et al., 2020). Automatic self-prioritization has also been demonstrated in endogenous attentional processes in working memory tasks (Yin et al., 2019). Self-prioritization with auditory stimuli may thus be diminished when auditory or audiovisual self-associations compete with (putatively stronger) visual self-associations in intermixed trials.

Alternatively, visual dominance arising in intermixed trials as a result of attentional effects could reduce the SPE with auditory stimuli. For example, the alerting properties of auditory stimuli are thought to elicit an override response in participants such that they actively focus their attention toward the weaker visual stimuli as a consequence. Fewer cognitive resources thus remain to attend to, and process, auditory stimuli (Posner et al., 1976). Stolte et al. (2021) examined the SPE across visual and audiovisual stimuli only in the matching task, so it is not known whether the SPE would be similarly reduced with unisensory auditory stimuli in the task (i.e., auditory-only object and label stimuli).

**Auditory stimulus intensity** The SPE has been shown to be affected by low-level sensory features of visual stimuli (i.e., stimulus contrast): namely, self-bias increased when stimulus contrast was reduced (Sui et al., 2012 – Experiment 4). Sui et al. (2012) suggested that this experimental manipulation could be considered a perceptual modulation. As noted, in contrast to Stolte et al. (2021), Schäfer et al. (2016b) presented their auditory stimuli at 50 dB rather than 75dB. Here it is worth bearing in mind that sound intensity has been shown to modulate modality-specific (perceptual) processes, as well as the motor output (Miller et al., 1999; St Germain et al., 2020), and also multisensory processes (Barutchu et al., 2010; Ma et al., 2009). If the SPE were to interact with stimulus intensity, it might be expected to

increase with auditory stimulation at lower as compared with higher intensities. In other words, if self-relevance in the matching task can boost auditory perceptual processes as has been proposed for visual perceptual processes, the SPE would increase if reduced auditory stimulus intensity is detrimental to stranger-associated, but not self-associated, responding.

**Modalities of the stimulus objects and labels** Typically, previous studies examining the SPE with auditory and visual stimuli have used visual labels (Payne et al., 2020; Schäfer et al., 2015; Schäfer et al., 2016b; Stolte et al., 2021). To date, no studies have examined the SPE with unisensory auditory stimuli (simultaneously presented auditory object and auditory label). That said, the self-bias arising with unisensory auditory stimuli (e.g., the participant's own name) has been demonstrated in other paradigms, for example, in dichotic listening tasks (e.g., Moray, 1959). Notably, however, the auditory self-bias that was first shown in Moray's early research was exhibited by only  $\approx 33\%$  of the participants (Conway et al., 2001; Wood & Cowan, 1995). Furthermore, differential mechanisms (with which the SPE could potentially interact) have been shown to underlie the processing of visual and auditory linguistic and object stimuli carrying seemingly equivalent information (Arana et al., 2020; Chen & Spence, 2018). The sensory modality of the label (auditory or visual) may therefore not be interchangeable, and the SPE may be moderated by the combination of object/label modality type. It could be that visual object-label self-associations are more easily formed, or accessed, than unisensory auditory or multisensory object-label self-associations. Alternatively, however, *unisensory* (rather than just visual) self-associations may be more easily formed or accessed than multisensory self-associations.

Attention has also been shown to modulate the SPE (Humphreys & Sui, 2016), and attentional effects arising in responses to auditory and audiovisual stimuli may also interact with the SPE. For example, while attentional-capture is likely to increase with higher-intensity sounds, when those sounds are presented simultaneously with visual stimuli (i.e., as a multisensory warning signal), such effects may be attenuated (Spence & Driver, 1999). In other words, the SPE may be moderated across unisensory auditory and audiovisual stimuli.

## The present study

The present study explored whether stimulus-related and task-design parameters would moderate the SPE in a matching task requiring motor responses to unisensory and multisensory stimuli. We also examined (for the first time) whether the SPE would transfer to a multisensory simple detection motor response paradigm whereby, in contrast to the matching task,

the associations of the stimuli with the self were irrelevant to the task at hand (cf. Orellana-Corrales et al., 2021; Stein et al., 2016; Wade & Vickery, 2018; Woźniak & Knoblich, 2021).

In order to compare the SPE in the matching task across unisensory auditory stimuli, as well as unisensory visual and multisensory stimuli, we introduced auditory labels (see Fig. 1). The use of auditory labels also ensured equal probability of occurrence of the visual and auditory stimuli across all trials. (In a version of the matching task using visual and auditory object stimuli, with only visual label stimuli, a visual stimulus but not an auditory stimulus would be presented in every trial).

Two independent groups of participants took part in the present study to compare the SPE across auditory stimulus intensities. The participants in Group 1 completed an audio-visual adaptation of the matching task (Sui et al., 2012), responding to unisensory auditory, unisensory visual, and multisensory object-label pairs, in both blocked and intermixed trials, and with the auditory stimuli presented at 50 dB (following Schäfer et al., 2016b). For the participants in Group 2, the procedure was identical except that the auditory stimuli in the matching task were presented at 70 dB (consistent with Stolte et al., 2021). The participants in this latter group then subsequently also completed a simple detection task in which they made motor responses to the shape and sound stimuli and their multisensory combinations. Previous studies have examined whether the self-associations formed in the matching task can transfer to other task paradigms (Chiarella et al., 2020; Dalmaso et al., 2019; Hu et al., 2020; Moradi et al., 2018; Payne et al., 2017; Stein et al., 2016;

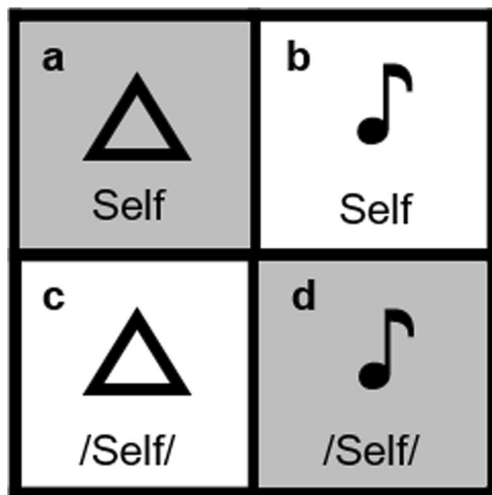
Woźniak & Knoblich, 2019; Yin et al., 2019). To the best of our knowledge, though, the present study is the first to examine whether the SPE can transfer across tasks to a simple multisensory detection motor response paradigm.

It was expected that the SPE in the matching task would be moderated by the block type (blocked or intermixed trials), and reduced with audiovisual as compared to visual stimuli (e.g., Stolte et al., 2021), although the latter may depend on task context (block type and auditory stimulus intensity; cf. Schäfer et al., 2016b). It was further hypothesized that while a modulation of the SPE across stimulus modality types might be expected (i.e., if the SPE interacts with modality-specific processes or differential stimulus effectiveness), a consistently reduced SPE with auditory stimuli would lend support to Stolte et al.'s findings. An equivalent SPE across stimulus modality types, on the other hand, would lend support to the contention that the SPE is underpinned by a modulation of modality-general processes (in line with Schäfer et al., 2016b). In addition, if an SPE arose in the detection task, this would further suggest that the unisensory and multisensory self-associations established in the matching task can be automatically activated in fast motor responses to the unisensory and multisensory stimuli.

## Methods

### Participants

**Matching tasks** In the present study, the SPE was examined using a three-way  $2 \times 4 \times 2$  mixed design: 2 between groups (Group 1 in which the auditory stimuli were presented at 50 dB, and Group 2 in which the auditory stimuli were presented at 70 dB), and, within each group, we examined the effects of the four sensory modality stimulus types (unisensory visual, unisensory auditory, multisensory with visual labels, multisensory with auditory labels),  $\times$  2 block types (trials blocked by sensory modality type, intermixed trials of the unisensory and multisensory stimuli). Multisensory gains in the matching task were examined using a 2 (Auditory stimulus intensity between-groups Groups 1 and 2: 50 dB, 70 dB)  $\times$  2 (within-groups Block type: blocked, intermixed)  $\times$  2 (within-groups AV Stimulus type: A+VL, V+AL)  $\times$  2 (within-groups Association: self, stranger) mixed factorial structure. Previous studies reported a large self-bias effect using the standard visual modality matching task ( $d_z > 0.80$ ; Sui et al., 2012) and a medium effect-size using an adaptation of the task using auditory objects and visual labels in a blocked-trial design ( $d_z = 0.51$ ; Schäfer et al., 2016b). Stolte et al. (2021) reported a large effect size for the interaction between association and stimulus type using an intermixed-trial design ( $\eta^2 = .15$ ; Stolte et al., 2021 – Table 1). For a medium effect size ( $\eta_p^2 = .12$ ), a probability of  $1 - \beta = 0.80$ , and an  $\alpha$ -value of 0.05,



**Fig. 1** Unisensory stimuli and multisensory stimuli used in the matching task. Example using the self-associated label. Simultaneously-presented: **a.** Unisensory visual stimulus type: Visual object + Visual Label = V+VL. **b.** Multisensory stimulus type: Auditory Object + Visual Label = A+VL. **c.** Multisensory stimulus type: Visual object + Auditory label = V+AL. **d.** Unisensory auditory stimulus type: Auditory object + Auditory label = A+AL. ♪ = categorical sound.  $\Delta$  = geometric shape (triangle presented as example). The Self label is depicted. /Self/ denotes the spoken self-associated label



a minimum sample size of 28 participants was required for a  $2 \times 4 \times 2$  between-within-within ANOVA and 60 for a  $2 \times 2 \times 2 \times 2$  between-within-within-within ANOVA (MorePower 6.0.4 program; Campbell & Thompson, 2012)

Sixty right-handed participants (14 male, ages 18–23 years, mean age  $18.92 \pm 1.00$ ) with self-reported normal or corrected-to-normal visual acuity and hearing took part in the study. In Group 1, there were 31 right-handed participants (seven male, ages 18–23 years, mean age  $19.03 \pm 1.02$ ); in Group 2, there were 29 right-handed participants (seven male, ages 18–22 years, mean age  $18.79 \pm 0.98$ ). The participants were recruited via the Oxford University Research Participation Scheme and received course credit for their time and effort. A written consent form approved by the University of Oxford Central University Research Ethics Committee (MS-IDREC- R61057/RE002) was completed by all participants.

**Detection task** No previous studies have looked at the SPE in a multisensory simple detection task, and so the effect size was unknown. Given that the effect size of the SPE using visual or auditory stimuli is typically medium to large (Schäfer et al., 2016b; Sui et al., 2012), a moderate effect size ( $f = 0.25$ ), a probability of  $1 - \beta = 0.80$ , and an  $\alpha$ -value of 0.05,

would require a minimum sample size of 16 participants. (G\*Power 3.1 program; Faul et al., 2009). In Group 2, 29 right-handed participants (seven male, ages 18–22 years, mean age  $18.79 \pm 0.98$ ) with normal or corrected-to-normal visual acuity and hearing completed the matching task. Twenty-six participants (seven male, ages 18–22 years, mean age  $18.73 \pm 1.00$ ) of the 29 who completed the matching task completed the detection task (three participants who completed the matching task did not complete the detection task).

**Stimuli and apparatus**

**Matching tasks** All computer tasks were conducted on a PC with a 23-in. LCD monitor ( $1,920 \times 1,080$  pixels at 60-Hz refresh rate) using E-Prime software (version 2.0). A QWERTY-keyboard recorded button-press responses. Following previous studies (Sui et al., 2012), participants made matching and mismatching responses using their index and middle finger of their right hand on two adjacent keyboard keys. Visual object stimuli consisted of two geometric shapes (V) from the following set (pentagon, hexagon, or octagon, each subtending  $3.2 \times 3.2^\circ$  of visual angle), and two written-text self- and stranger-related labels (VL) (*your*, *their*, referring to ‘your shape’, ‘their shape’), subtending a visual angle

**Table 1** Mean RT, Accuracy (percentage correct), and Sensitivity (D-Prime;  $d'$ ) index scores with Standard Deviations, as a Function of Stimulus Type (Visual-shape+Visual-Label, A+AL, A+VL, V+AL),

Association (self, stranger), Block Type (blocked, intermixed), and Match Condition (match, mismatch) in the Matching tasks

Stim. type	Assoc.	RT				Percentage correct				$d'$	
		Blocked		Intermixed		Blocked		Intermixed		Block	Intermixed
		Match	Mismatch	Match	Mismatch	Match	Mismatch	Match	Mismatch		
Group 1 (50dB) (N = 28)											
V+VL	Self	674 (57)	752 (68)	720 (63)	808 (62)	92 (9)	81 (14)	86 (10)	82 (13)	2.65 (0.93)	2.27 (0.87)
	Stranger	735 (76)	763 (62)	771 (71)	791 (57)	80 (16)	84 (11)	81 (14)	83 (11)	2.20 (1.05)	2.06 (0.95)
A+AL	Self	714 (88)	791 (69)	805 (60)	862 (48)	86 (12)	77 (18)	79 (17)	71 (17)	2.12 (1.11)	1.55 (1.02)
	Stranger	757 (64)	764 (56)	853 (61)	844 (53)	79 (17)	83 (15)	63 (17)	74 (18)	2.02 (1.05)	1.12 (1.06)
A+VL	Self	673 (88)	744 (59)	776 (63)	835 (60)	91 (11)	87 (12)	85 (16)	77 (19)	2.87 (1.08)	2.08 (1.19)
	Stranger	732 (71)	721 (56)	812 (71)	814 (57)	86 (11)	89 (9)	75 (16)	78 (17)	2.69 (0.89)	1.71 (1.06)
V+AL	Self	627 (83)	717 (55)	719 (57)	790 (56)	92 (9)	90 (10)	87 (10)	86 (11)	3.00 (0.83)	2.49 (0.83)
	Stranger	720 (78)	715 (58)	780 (70)	772 (53)	86 (10)	93 (6)	81 (14)	87 (11)	2.91 (0.97)	2.20 (0.89)
Group 2 (70dB) (N = 22)											
V+VL	Self	664 (75)	769 (71)	700 (46)	784 (56)	90 (15)	80 (11)	87(10)	80 (19)	2.48 (0.94)	2.22 (1.00)
	Stranger	716 (82)	758 (64)	745 (68)	775 (54)	80 (17)	85 (14)	83 (15)	82 (15)	2.38 (1.35)	2.23 (1.21)
A+AL	Self	714 (79)	794 (55)	809 (84)	858 (73)	87 (10)	81 (17)	73 (17)	65 (20)	2.38 (1.24)	1.18 (1.03)
	Stranger	746 (105)	769 (71)	811 (96)	841 (67)	78 (20)	82 (16)	65 (23)	70 (22)	2.06 (1.34)	1.15 (1.34)
A+VL	Self	683 (71)	781 (78)	777 (68)	836 (71)	93 (7)	84 (14)	86 (10)	72 (17)	2.76 (1.09)	1.84 (0.84)
	Stranger	735 (96)	746 (65)	801 (94)	819 (64)	82 (13)	89 (9)	73 (19)	80 (15)	2.41 (0.84)	1.72 (1.09)
V+AL	Self	628 (79)	726 (97)	713 (64)	771 (73)	95 (5)	86 (15)	89 (9)	84 (19)	3.05 (0.99)	2.58 (1.05)
	Stranger	696 (111)	711 (101)	744 (73)	763 (75)	87 (12)	90 (9)	82 (16)	88 (16)	2.77 (1.09)	2.51 (1.28)

Note. RT Reaction time, Standard deviations appear within parentheses. Stim. type Stimulus type, Assoc. Association

of  $2.1 \times 0.7^\circ$ . The words ‘Your’ and ‘Their’ were chosen as labels due to their similar word length and equivalent number of syllables (durations), and equivalent low ratings for ‘word concreteness’ (Brysbaert et al., 2014; see Desebrock & Spence, 2021). The two shapes allocated to each participant and the labels that they were paired with were counterbalanced across participants following a Latin square design. The same shape and label pairs were maintained for the blocked and intermixed presentations. The shapes and labels were presented against a black background in the centre of the PC-screen. The shape was positioned above (and the label below) a fixation cross ( $0.6 \times 0.6^\circ$  of visual angle).

Auditory object stimuli (A) consisted of two neutral instrumental sound samples (150-ms duration) from the following set (violin, synth vocal, and clarinet). The neutral instrumental sound samples were selected from two validated, publicly available sound sets: the Musical Emotional Bursts (MEB; Paquette et al., 2013) and the Montreal Affective Voices (MAV; Belin et al., 2008). The MEB is a set of musical affect bursts expressing basic emotional states (specifically, happiness, sadness, and fear) and also ‘neutral’ expressions. Two neutral sounds from the MEB and one neutral sound sample from the MAV on the same musical note were selected for the present study: “V3\_NEUTRAL\_MEB” (violin), “45\_NEUTRAL\_MAV” (vocal), and “C1\_NEUTRAL\_MEB” (clarinet). All sound samples were trimmed to a duration of 150 ms. The auditory labels (AL) were two recorded samples (bursts) of fast natural speech utterances of the words “your” and “their” spoken by a female actor. The auditory stimuli were presented via two free-standing loudspeakers, one placed on either side of the PC screen at a distance of 50 cm from the participant. Schäfer et al. (2016b) used auditory stimuli of low intensity (50 dB) just above ambient sound levels in their sound and visual label version of the matching task. In the present study the sound pressure level (SPL) of the ambient background noise was set at 45 dB. The auditory stimuli were presented at 50 dB SPL (following Schäfer et al., 2016b) to one group of participants (Group 1;  $N = 31$ ) and at 70 dB (consistent with Stolte et al., 2021) for a second group (Group 2;  $N = 29$ ). Signal-to-noise (SNR) ratios were thus 5 dB and 25 dB for the two groups, respectively. Auditory and visual object-label stimuli were combined to produce four stimulus types: visual only (object and label both visual = V+VL), auditory only (sound and label both auditory = A+AL), auditory with visual label (A+VL), and visual object with auditory label (V+AL). All of the stimuli were presented simultaneously for 150 ms.

**Detection task** A Cedrus RB-530 response-box recorded button-presses in the simple detection task. (A response box was used for this task due to the robust nature, size, and spacing of the response-box buttons which could accommodate

the fast motor responses required by the detection task – the response box was only used for the simple detection task.) The stimuli consisted of the two self- and stranger-associated geometric shapes and instrumental sounds that the participant had been assigned in the preceding matching tasks. (NB: The visual and auditory labels used in the matching task were not used as stimuli in the detection task.)

## Procedure

**Matching tasks** Participants carried out an extended multisensory version of Sui et al.’s (2012) computer-based matching task, in a single experimental testing session. The duration of the testing session ranged from 90 to 120 min. The matching task consisted of four blocks of blocked trials (one stimulus type per block; 80 trials per block), the order of which was counterbalanced across participants following a balanced Latin square design, followed by five blocks of randomly intermixed trials (all four stimulus types presented randomly with equal probability within each block; 96 trials per block). Familiarity with both the matching task and the 16 stimulus-response mappings ( $2 \times \text{Association} \times 2 \text{ Match-type} \times 4 \text{ stimulus types}$ ) gained during the blocked trials enabled participants to complete the intermixed trials, which were otherwise too difficult. (See Appendix 6 for further details regarding task order.)

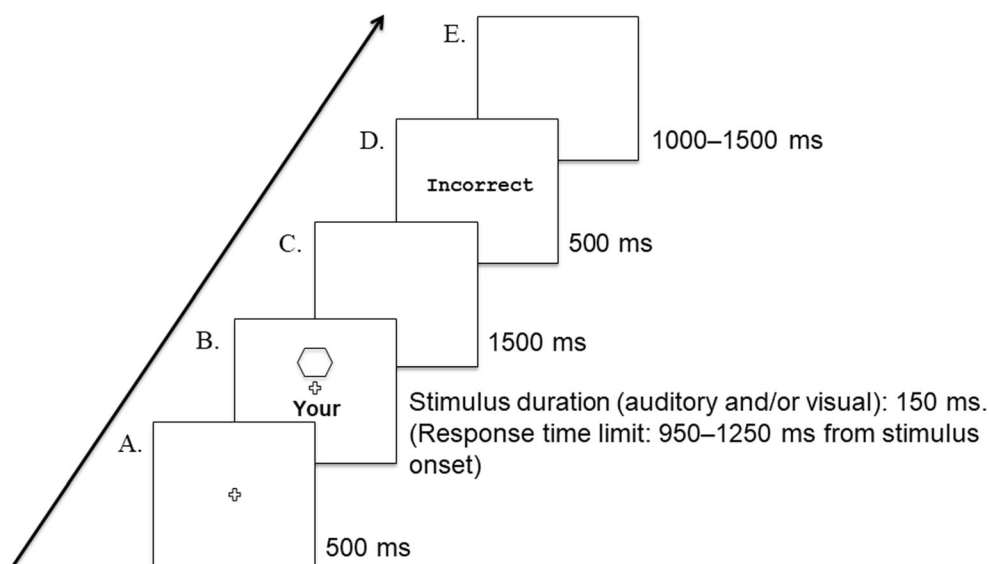
Following Sui et al.’s (2012) procedure, for the visual stimulus blocks (V+VL), participants were instructed (via onscreen text) to associate one of their allocated geometric shapes or sound with themselves (specifically, as ‘your’ shape; e.g., ‘the pentagon is your shape’) and the second allocated shape or sound with ‘a stranger’ (as ‘theirs’; e.g., ‘the octagon is their shape’). Order of instructions pertaining to ‘self’ and ‘stranger’ associations were counterbalanced across participants. Following this, the participants carried out the matching task in which participants used their right hand to depress one of two response keys (b or v) on the keyboard to indicate either a ‘matching’ or a ‘mismatching’ judgment (the keys and match/mismatch mappings were counterbalanced across participants). The participants were instructed to make their responses to the stimuli as rapidly and accurately as possible. The main task was preceded by 24 practice trials, with a performance accuracy threshold set at 60% correct (that is, participants had to achieve at least 60% correct before they could proceed to the main task). Onscreen feedback was presented (*Correct*, *Incorrect*, *Too Slow*) during the practice trials, with the ‘*Correct*’ feedback omitted during the main task (following Schäfer et al., 2016b).

The procedure was the same for all blocks except that in the A+VL block, participants matched instrumental sounds with visual labels (e.g., clarinet-‘your’, violin-‘their’), in the V+AL block they matched shapes with auditory labels, in the A+AL block they matched the instrumental sounds with the auditory

labels. In intermixed blocks, the participants matched the visual and auditory labels with the shapes and sounds that they had been allocated in the previous blocks, with all stimulus modality types intermixed within each block of trials. Prior to carrying out the intermixed-trial blocks, the participants completed 32 practice trials. There was an 8-s break between each block of trials. There were 800 trials in total (320 blocked trials, and 480 intermixed trials), across 32 conditions – 2 block types (blocked/intermixed)  $\times$  4 stimulus types (V+VL, A+AL, A+VL, V+AL)  $\times$  2 associations (self/stranger)  $\times$  2 matching types (match/mismatch), 20 trials per condition in blocked trials, and 30 trials per condition in the intermixed trials. Each condition was randomly generated with an equal number of presentations. The participants were informed of their overall accuracy at the end of each block of trials. A schematic representation of an experimental trial in the matching task is shown in Fig. 2

**Detection task** After completing the matching task, the participants in Group 2 completed a simple detection task in which they were asked to respond as rapidly and accurately as possible to the auditory (instrumental sounds), visual (shapes), and audio-visual (shape-sound) stimuli used in the matching task. In a modified version of the typical detection task, a short response time limit of 300 ms was used to ensure that participants did not delay their responses (to avoid the intentional evaluation of the stimuli), and to keep participants alert to the task, responding as rapidly as they could ( $< 3\%$  of reaction times (RTs) were excluded based on this response limitation). Only the shape and sound stimuli were used (the labels were not used). As such, the stimuli pairs in the detection task

differed from those in the matching tasks (which used labels with individual shape/sound stimuli). The unisensory auditory (self- or stranger-associated), unisensory visual (self- or stranger-associated), and audio-visual stimuli (self-match, stranger-match, stranger-mismatch, and self-mismatch) were presented in a random order across four blocks of 80 trials. Self-match trials consisted of the self-associated shape and self-associated sound stimuli, and stranger-match trials consisted of the stranger-associated shape and stranger-associated sound stimuli. The stranger mismatch trials consisted of the stranger-associated shape and self-associated sound stimuli, while the self-mismatch trials consisted of the self-associated shape and stranger-associated sound stimuli. There was a total of eight conditions: A self-associated unisensory auditory stimulus type, a stranger-associated unisensory auditory stimulus type, a self-associated unisensory visual stimulus type, a stranger-associated unisensory visual stimulus type, two self-associated audiovisual stimulus types, and two stranger-associated audiovisual stimulus types. There were 40 trials per condition. Sixteen practice trials preceded the main blocks with a performance threshold set at 80% accuracy – participants could proceed to the main blocks once their performance (responding within 300 ms of stimulus onset) reached or exceeded 80% accuracy. To discourage anticipatory responses, “Too early” feedback was presented if the participant responded during the presentation of the fixation cross and “too slow” was displayed if they responded outside of the time limit, or not at all (such responses were also recorded as incorrect and excluded from the analysis). If participants responded correctly within the time-limit, no feedback was presented. (After the computer



**Fig. 2** Schematic overview of an experimental trial sequence for the matching task. (Displayed elements not to scale). **a** Fixation cross. **b** Visual, auditory, or audiovisual stimulus onset. Shape-label (*Visual Task*) stimulus shown. **c** Blank screen. **d** Written feedback displayed on

screen – “Incorrect” / “Too slow” (for a ‘correct’ response a blank slide was displayed following Schäfer et al., 2016b). **e** Inter-trial intervals generated at random

tasks, a sub-group of participants completed questionnaires measuring individual differences on self- and other-related dimensions (e.g., ‘personal distance’; Sui & Humphreys, 2015). The data from these instruments will be analysed and presented as part of a separate future study.

## Data analysis

**Matching tasks** There were two main output measures: RTs (measured from stimulus onset to the depression of the keyboard key), and percentage of correct responses (accuracy). Following previous research (Sui et al., 2012), a signal detection approach was used to calculate an index of sensitivity ( $D'$ ; Green & Swets, 1966). Hits were coded as *yes* responses to match trials, and false alarms were coded as *yes* responses to mismatch trials with the same shape. Mismatch conditions were defined as either shape- or sound-based (i.e., a self-mismatch trial consisted of the self-associated shape or sound and the stranger-associated label, a stranger-mismatch trial consisted of the stranger-associated shape or sound and the self-associated label). RTs were based on correct responses. RTs above or below 2.5 SDs from individual means were trimmed (less than ~1% of RTs were excluded). For absolute measures of RT, percentage accuracy, and  $d'$  data, see Table 1.

Self-bias index scores were calculated using RTs and  $d'$ . Normalized self-bias scores in RTs were calculated using the matching-condition RTs (e.g., Constable et al., 2021; Desebrock & Spence, 2021; Sui & Humphreys, 2017a). Match-trial stimuli (e.g., a self-associated object and self-associated label) involve one association, thus in behavioural paradigms where effects in mismatch trials cannot be disentangled, matching-trial responses index self- and stranger-related processing (Sui et al., 2012). Mismatch trials are typically treated as fillers in the literature on the SPE (see Desebrock & Spence, 2021; Schäfer et al., 2016b). Therefore, in accordance with the aims of the present study, and following the rationale of previous research, the focus of the analysis reported here was on match-trial RT data. Self-bias scores were given by the formula: “(stranger – self)/(stranger + self)” for RTs. For  $d'$ , self-bias was indexed by the differential scores (self – stranger) between self-associated and stranger-associated conditions (following Sui & Humphreys, 2017a). Positive values indicate an advantage for self. (NB: Since auditory stimulus intensity influences RTs and  $d'$ , main effects of dB level on RT and  $d'$  across Groups 1 and 2 were assessed. There was no main effect of auditory stimulus intensity on RTs or  $d'$  (see Appendix 2), therefore  $d'$  self-bias index scores were not normalised.)

Normalised differential scores – self-bias index scores (e.g., Constable et al., 2021; Desebrock & Spence, 2021; Schäfer et al., 2016b; Sui & Humphreys, 2017a) – were analysed to examine whether the magnitude of the self-

advantage in the matching task responses was modulated by stimulus type, block type, and auditory stimulus intensity. These scores provide an index of the relative magnitude of the difference in performance between self- and stranger-related responses. The present study follows previous studies that have examined contextual factors (both stimulus- and task-design-related) that moderate the well-established self-advantage in the matching task using different task parameters (e.g., Golubickis et al., 2017; Golubickis & Macrae, 2021; Hu et al., 2020; Stolte et al., 2021; Verma et al., 2021; Woźniak & Knoblich, 2019). The hypotheses of the present study thus spoke to the modulation rather than the emergence of the SPE in the matching paradigm, so the focus was on analysing differences in the magnitude of the SPE rather than absolute RTs (e.g., see Constable et al., 2021). See Appendix 1 for the analyses of absolute RTs and  $d'$  measures. A significant advantage for self-associated responses was found in both the Group 1 ( $\eta_p^2 = .62$ ) and Group 2 samples ( $\eta_p^2 = .23$ ) (Appendix 1).

To further assess the self-advantage in responses to unisensory as compared with multisensory stimuli in the matching task, multisensory gains/costs in RTs and sensitivity ( $d'$ ) in the self- and stranger-associated matching conditions were compared. Multisensory facilitation in both the blocked and the intermixed trials in the matching tasks was assessed in RTs and  $d'$  by subtracting the RTs/ $d'$  scores of responses in the multisensory conditions from the fastest/highest of the counterpart unisensory conditions (e.g., Barutçu et al., 2018, 2020). For example, multisensory facilitation in RT for self-associated responses to the A+VL stimulus type was calculated by deducting A+VL self-associated matching-trial RTs from the fastest of the V+VL and A+AL self-associated matching-trial RTs. Multisensory facilitation in RT for self-associated responses to the V+AL stimulus type was calculated by deducting V+AL self-associated matching-trial RTs from the highest of the V+VL and A+AL self-associated matching-trial RTs. Multisensory facilitation in  $d'$  scores for self-associated responses to the A+VL stimulus type was calculated by deducting A+VL self-associated condition  $d'$  scores from the highest of the V+VL and A+AL self-associated condition  $d'$  scores (and then multiplying by -1 so that positive values indicated gains). Multisensory facilitation in  $d'$  scores for self-associated responses to the V+AL stimulus type was calculated by deducting V+AL self-associated condition  $d'$  scores from the highest of the V+VL and A+AL self-associated condition  $d'$  scores (and then multiplying by -1 so that positive values indicated gains). Thus, the multisensory gain value was the difference between the multisensory and unisensory responses. A positive value indicated faster RTs/higher  $d'$  and a negative value indicated slower RTs/lower  $d'$  (multisensory costs) for multisensory self- or stranger-associated responses as compared with the fastest/highest unisensory self- or stranger-associated responses. (NB: Mean SPE gains were also calculated but not analysed; they are



depicted in Fig. 5c. SPE gains were calculated by deducting the self-associated multisensory condition – Matching- or Mismatching-trial – RTs from the stranger-associated multisensory condition RTs. Positive values indicate a self-advantage.) In order to compare multisensory gains/costs in RTs across Group 1 and Group 2, the multisensory gain/cost measures were converted into percentage gains/costs to control for differences in processing speed, and were given by the formula: [(faster of unisensory – multisensory)/faster of unisensory] × 100.

**Detection task** RTs < 100 ms and > 2.5 SDs above individual means were trimmed, excluding < 0.5% of total RTs. Percentage detection rates, RTs, and central tendencies of the RT distributions were examined. There were two within-participant factors: Association on two levels (self, stranger) and Condition with four levels (Auditory, Visual, AV Mismatch, AV Match). Cumulative Distribution Frequency plots (CDFs; Ratcliff, 1979) were also constructed using CDF-XL (Houghton & Grange, 2011) to further assess distributional information and examine differences throughout the whole RT distributions. Percentiles (deciles) for the rank-ordered RTs by condition, for each participant, were calculated. For each condition, cumulative probabilities were calculated from 0.1 up to 1.0 in increments of 0.10. The CDF plots were then drawn using MATLAB. To preserve power, every second probability was included in the analyses. There were three within-participant factors: Association on two levels (self, stranger), Condition on four levels (A, V, AV Mismatch, and AV Match), and Probability on five levels (0.1, 0.3, 0.5, 0.7, 0.9).

Central tendencies were similarly calculated for MS gains (calculated by subtracting the RTs of responses in the multisensory conditions from the fastest of the counterpart unisensory conditions). There were two within-participant factors: Association on two levels (self, stranger) and AV Condition on two levels (Match, Mismatch).

Effect sizes were calculated using Cohen's  $d_z$  for t-tests and partial eta-squared ( $\eta_p^2$ ) for ANOVAs (Cohen, 1988; Lakens, 2013). To adjust for multiple comparisons, Holm-Bonferroni corrections at an  $\alpha$  value of .05 were applied (Holm, 1979) with unadjusted significance values reported. For violations of sphericity, Greenhouse-Geisser correction were applied where appropriate.

## Results

### Matching tasks

In the 50 dB group (Group 1), the data from two participants were excluded for having chance accuracy ( $M < 50\%$ ) and (< 30% correct in more than one condition). The data were assessed for outliers in the match-trial and mismatch-trial RT

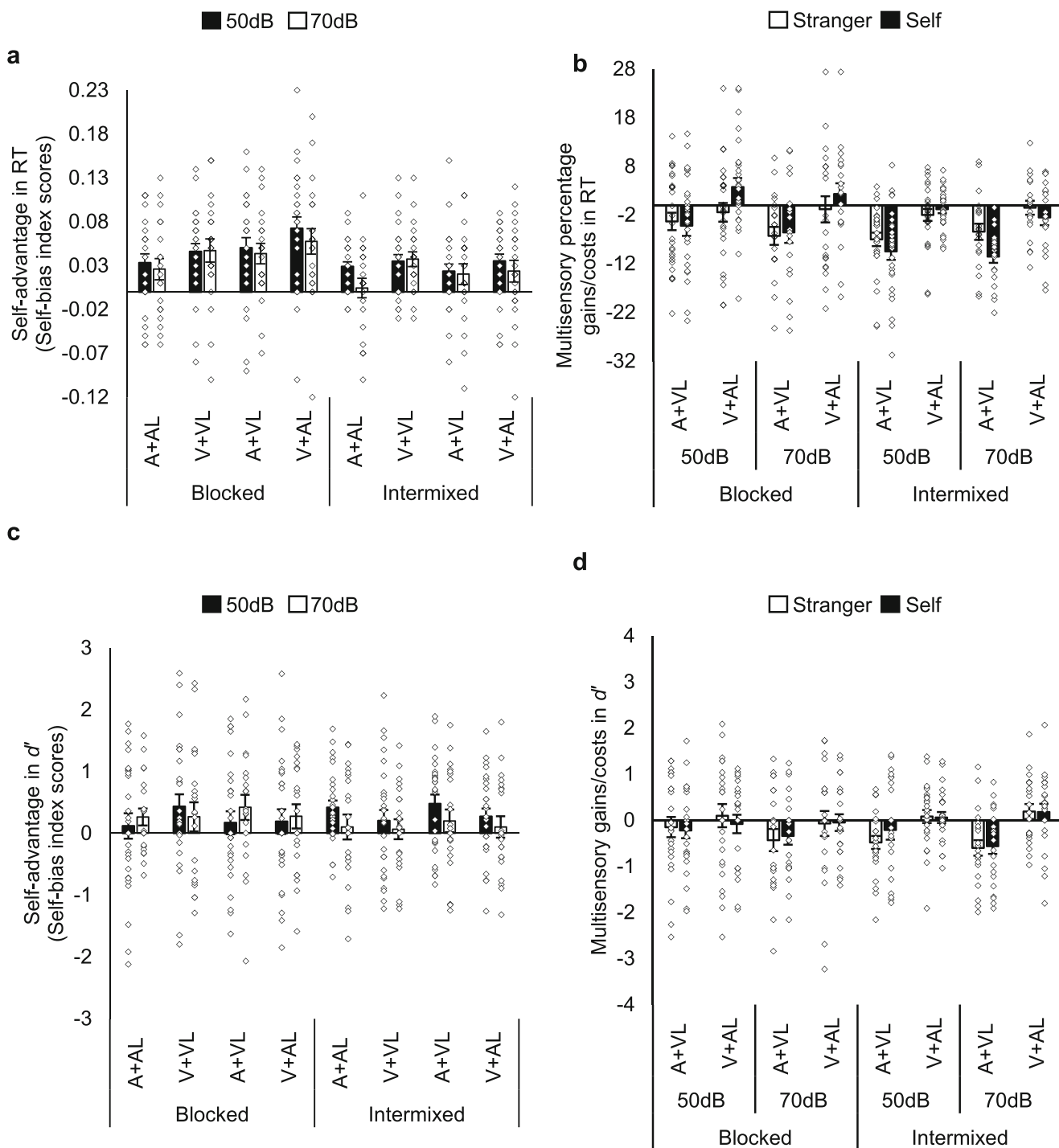
data, and the  $d'$  data (studentized residuals outside  $\mp 3.0$  in absolute value in one or more conditions), and one participant's data was identified as an outlier and thus removed from the analysis (mean RTs and  $d'$  data are presented in Table 1). In the 70-dB group (Group 2), five participants were excluded for having chance levels of accuracy, and the data from two participants were excluded as they constituted outliers. The data from a total of 50 participants (28 in Group 1 and 22 in Group 2) were therefore used in the analysis. Self-bias index scores for RT and  $d'$ , and multisensory RT,  $d'$ , gains and costs, are presented in Fig. 3

### Self-bias in reaction times (RTs)

The self-bias in RT data is presented in Fig. 3a. There were two outliers (studentized residuals outside  $\mp 3.0$  in absolute value). Both outliers were removed ( $N = 48$ ), and since the majority of the data were normally distributed (two conditions not normally distributed), we chose not to transform the data. The assumption of homogeneity of variances was violated for one condition, as assessed by Levene's test for equality of variances. Given that the sample sizes of the two groups were roughly equal (ratio 1.29) and ANOVA is generally robust to violations of this assumption if group sizes are roughly equal, ANOVA was carried out on the data. A 2 (Auditory stimulus intensity: 50 dB, 70 dB) × 2 (Block type: Blocked, Intermixed) × 4 (Stimulus type: V+VL, A+AL, A+VL, V+AL) mixed ANOVA was conducted on the RT self-bias index scores. The analysis revealed main effects of Stimulus type,  $F(2.21, 101.47) = 4.76$ ,  $p = .009$ ,  $\eta_p^2 = .09$ , and Block type,  $F(1, 46) = 24.74$ ,  $p < .001$ ,  $\eta_p^2 = .35$ . Overall, the magnitude of self-bias in the Blocked trials ( $M = .05$ ,  $SE = .006$ ) was greater than in the Intermixed trials ( $M = .03$ ,  $SE = .005$ ). For Stimulus type, the magnitude of the normalised self-bias in the RT data was significantly greater ( $p = .002$ ) in responses to V+AL ( $M = .05$ ,  $SE = .007$ ) as compared with A+AL ( $M = .02$ ,  $SE = .006$ ) (see Fig. 4). There were no significant differences between any other pair of Stimulus types ( $ps > .01$ ). There were no other significant effects. (The analysis was also conducted with the outliers included – see Table A1, Appendix 3. The findings were replicated.)

### Self-bias in sensitivity index scores (D-prime; $d'$ )

The self-bias in  $d'$  data is presented in Fig. 3b. There were two studentized residuals outliers. Both outliers were removed. The assumption of homogeneity of variances was violated for two conditions. As with the analysis of self-bias in the RT data, given that the group sizes were roughly equal, an ANOVA was carried out on the data ( $N = 48$ ). A 2 (Auditory stimulus intensity: 50 dB, 70 dB) × 2 (Block type: Blocked, Intermixed) × 4 (Stimulus type: V+VL, A+AL, A+VL, V+AL) mixed ANOVA was conducted on self-bias index scores



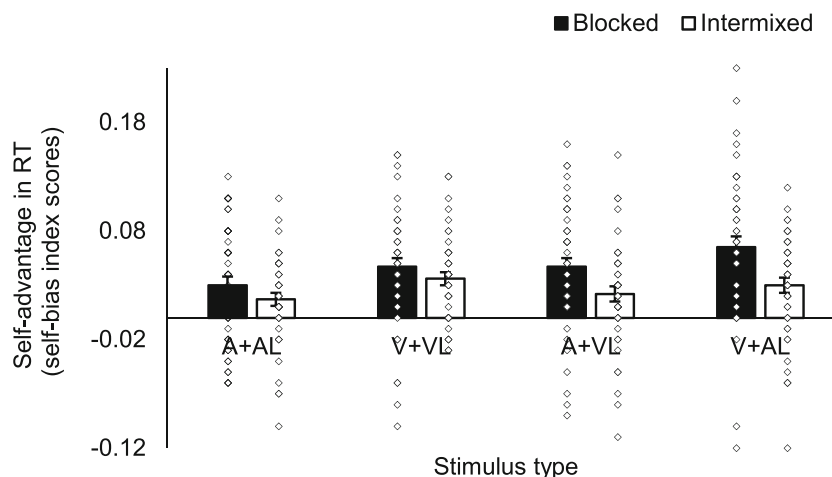
**Fig. 3** Bar graphs with individual data points (outliers excluded). Error bars represent SE. **a** Selfbias in RT index scores in the shape-label Matching trials as a function of Block type (blocked vs intermixed), Stimulus type, and Auditory stimulus intensity. **b** Multisensory percentage gains/costs in RT as a function of Block type, AV stimulus type, Association, and Auditory stimulus intensity. **c** Selfbias in sensitivity

index scores ( $d'$ ; D-prime) as a function of Block type, Stimulus type, and Auditory stimulus intensity. **d.** Multisensory gains/costs in  $d'$  as a function of Block type, AV stimulus type, Association, and Auditory stimulus intensity. V = visual shape stimulus, A = auditory stimulus, VL = visual (text) label, AL = auditory (spoken) label

for  $d'$ . There were no significant effects. The self-advantage in sensitivity did not differ significantly across the conditions. (The analysis was also conducted with the outliers included – see Table A1, Appendix 3. The findings were replicated.)

**Multisensory percentage gains/costs in RT**

The data are presented in Fig. 3c. There were two studentized residuals outliers whose removal produced a third outlier.



**Fig. 4** Bar graphs with individual data points (outliers excluded). Estimated marginal means of self-bias in RT index scores in shape-label Matching trials as a function of Block type and Stimulus type (with the Auditory stimulus intensity condition collapsed) showing the main

effects of Block type and Stimulus type, and the borderline-significant interaction between Block type and Stimulus type). Error bars represent SE. V = visual shape stimulus, A = sound stimulus, VL = visual (text) label, AL = auditory (spoken) label

With the three outliers removed, the data were submitted to a 2 (Auditory stimulus intensity: 50 dB, 70 dB)  $\times$  2 (Block type: blocked, intermixed)  $\times$  2 (AV Stimulus type: A+VL, V+AL)  $\times$  2 (Association: self, stranger) mixed ANOVA. There was a significant main effect of Block type,  $F(1, 45) = 8.85$ ,  $p = .005$ ,  $\eta_p^2 = .16$ , and AV Stimulus type,  $F(1, 45) = 50.14$ ,  $p < .001$ ,  $\eta_p^2 = .53$ . Negative gain values indicated that there were costs for responses to multisensory relative to unisensory stimuli. Multisensory costs were greater in participants' responses to intermixed trials ( $M = -4.75$ ,  $SE = 0.66$ ) as compared with blocked trials ( $M = -1.96$ ,  $SE = 1.05$ ), and greater in responses to the A+VL stimulus type ( $M = -6.34$ ,  $SE = 0.88$ ) as compared with responses to the V+AL stimulus type ( $M = -1.37$ ,  $SE = 0.83$ ). There was a significant interaction between Block type and Association,  $F(1, 45) = 8.04$ ,  $p = .007$ ,  $\eta_p^2 = .15$ , and between AV Stimulus type and Association,  $F(1, 45) = 5.07$ ,  $p = .03$ ,  $\eta_p^2 = .10$ . None of the other effects was significant. The interaction between the AV stimulus type and Association was probed revealing a significant difference between stranger-associated responses to the A+VL ( $M = -5.38$ ,  $SE = 1.02$ ) and V+AL ( $M = -1.44$ ,  $SE = 0.99$ ) stimulus types,  $p = .003$ , and a significant difference between self-associated responses to the A+VL ( $M = -7.30$ ,  $SE = 1.13$ ) and V+AL stimulus types ( $M = 0.70$  ms,  $SE = 0.97$ ),  $p < .001$ . There was no significant difference between self- and stranger-associated responses to the A+VL stimulus type,  $p = .13$ , or to the V+AL stimulus type,  $p = .05$ . However, it is worth noting that for A+VL stimuli, costs were descriptively greater for the self-associated than for the stranger-associated responses, while for the V+AL stimuli, costs were observed for the stranger-associated responses, and gains were uniquely observed for self-associated responses. Probing the Block type and Association interaction revealed greater multisensory costs for self-associated responses ( $M = -5.78$ ,  $SE = 0.80$ ) than

for stranger-associated responses ( $M = -3.71$ ,  $SE = 0.78$ ) in intermixed trials,  $p = .02$ . There was no significant difference between self-associated responses ( $M = -0.82$ ,  $SE = 1.24$ ) as compared with stranger-associated responses ( $M = -3.11$ ,  $SE = 1.18$ ) in blocked trials,  $p = .07$ . However, costs were descriptively greater in stranger-associated as compared with self-associated responses. In addition, stranger-associated responses in blocked trials ( $M = -3.11$ ,  $SE = 1.18$ ) as compared with intermixed trials ( $M = -3.71$ ,  $SE = 0.78$ ) were not significantly different,  $p = .63$ . In contrast, multisensory costs in self- as compared with stranger-associated responses were differentially modulated by block type. (The analysis was also conducted with the outliers included – see Table A2, Appendix 3. The findings were replicated except that there was no significant difference between self- and stranger-associated responses in intermixed trials,  $p = .06$ .)

#### Multisensory gains/costs in sensitivity index scores (D-prime; $d'$ )

The data for multisensory gains/costs in  $d'$  are presented in Fig. 3d. There were four studentized residuals outliers. With the outliers excluded, the data were submitted to a 2 (Auditory stimulus intensity: 50 dB, 70 dB)  $\times$  2 (Block type: blocked, intermixed)  $\times$  2 (AV stimulus type: A+VL, V+AL)  $\times$  2 (Association: self, stranger) mixed ANOVA. The analysis revealed a significant main effect of AV Stimulus type,  $F(1, 44) = 18.30$ ,  $p < .001$ ,  $\eta_p^2 = .29$ . Negative gain values indicated that there were costs for responses to multisensory relative to unisensory stimuli. There was some multisensory facilitation in responses to the V+AL stimuli ( $M = 0.06$ ,  $SE = 0.08$ ), and

costs in responses to A+VL stimuli ( $M = -0.36$ ,  $SE = 0.08$ ). There was a significant interaction between Block type and AV stimulus type,  $F(1, 44) = 6.21$ ,  $p = .02$ ,  $\eta_p^2 = .12$ . There were no other significant effects. The interaction between Block type and AV stimulus type was probed, revealing that costs in responses to the A+VL stimulus type ( $M = -0.28$ ,  $SE = 0.13$ ) as compared with the V+AL stimulus type ( $M = -0.02$ ,  $SE = 0.13$ ) in blocked trials were not significantly different ( $p = .05$ ). There were multisensory gains in responses to the V+AL stimulus type ( $M = 0.14$ ,  $SE = 0.08$ ) as compared with multisensory costs in responses to the A+VL stimulus type ( $M = -0.45$ ,  $SE = 0.09$ ) in the intermixed trials ( $p < .001$ ). (The analysis was also conducted with the outliers included – see Table A2, Appendix 3. The findings were replicated, except that there was a significant difference between multisensory costs in responses to the A+VL and V+AL stimulus types.)

### Detection task

We then examined, for the first time, whether the SPE can transfer to a simple detection task whereby, in contrast to the matching task, the self-associations were irrelevant to the task at hand (Orellana-Corrales et al., 2021; Stein et al., 2016; Woźniak & Knoblich, 2021). After completing the matching task, the participants in Group 2 made motor RT responses (single keypress) irrespective of the stimuli presented (Hecht et al., 2008; Miller, 1982; Wundt, 1910), which consisted of unisensory and multisensory combinations of the self- and stranger-associated auditory objects and visual shapes allocated to the participant in the matching task (without the labels). Completing the blocked and then intermixed trials of the matching task ensured that the participants were equally familiar with all stimulus modality types before completing the detection task.

The data from four participants were excluded for achieving  $> 2.5$  SDs below the group percentage detection rate mean for any one condition, or, for achieving  $< 50\%$  accuracy in any one condition. The data from 22 participants were used in the analysis. Percentage detection rate data, RT data, CDFs of RTs, and multisensory RT gains are all presented in Fig. 5

### Accuracy

For the detection rate data, see Fig. 5a. The studentized residuals for one condition were not normally distributed. Since the majority of the data were normally distributed, we chose not to transform the data. Percentage detection rate scores were submitted to a 2 (Association: Self, Stranger)  $\times$  4 (Stimulus type: Visual, Auditory, AV Mismatch, AV Match) repeated-measures ANOVA. There was a significant main effect of Stimulus type,  $F(3, 63) = 50.15$ ,  $p < .001$ ,  $\eta_p^2 = .71$ . There were no other significant effects. Pairwise comparisons between stimulus types revealed that detection rates in matching

**Fig. 5** Detection task.  $N = 22$  (Group 2). Bar graphs (a-c) with individual data points (outliers excluded). Error bars represent SE. **a** Mean Percentage Accuracy (detection rate) as a function of Association and Stimulus type. **b** Mean RTs as a function of Association and Condition (Stimulus type). **c** Mean multisensory RT gains in ms as a function of Association and Multisensory Condition (left). Mean SPE RT gains in ms as a function of multisensory condition (right). SPE gains = the difference between self- and stranger-associated responses. Positive values indicate a self-advantage. **d** Cumulative Distribution Functions (CDFs) of percentiles of the rank-ordered RTs (PC grand RT means) as a function of Condition (uni- and multisensory). AV Match self = the sound and shape associated with the self in the preceding matching task. AV Match stranger = the sound and shape associated with the stranger in the preceding matching task. MM = mismatch. AV self MM = self-associated shape and stranger-associated sound used in the preceding matching task. AV MM stranger = stranger-associated shape and self-associated sound used in the preceding matching task

( $M = 89.77$ ,  $SE = 1.27$ ) and mismatching ( $M = 89.77$ ,  $SE = 1.63$ ) multisensory trials were significantly greater than in unisensory visual ( $M = 73.35$ ,  $SE = 2.23$ ) and unisensory auditory ( $M = 77.39$ ,  $SE = 2.53$ ) trials ( $ps < .001$ ). There was no significant difference between AV Mismatch and AV Match Stimulus types ( $p > .99$ ), or between Auditory and Visual conditions ( $p = .05$ ; unadjusted significance value reported, Holm-Bonferroni correction, see *Data analysis*). Detection rates on AV trials were significantly higher than detection rates in unisensory trials.

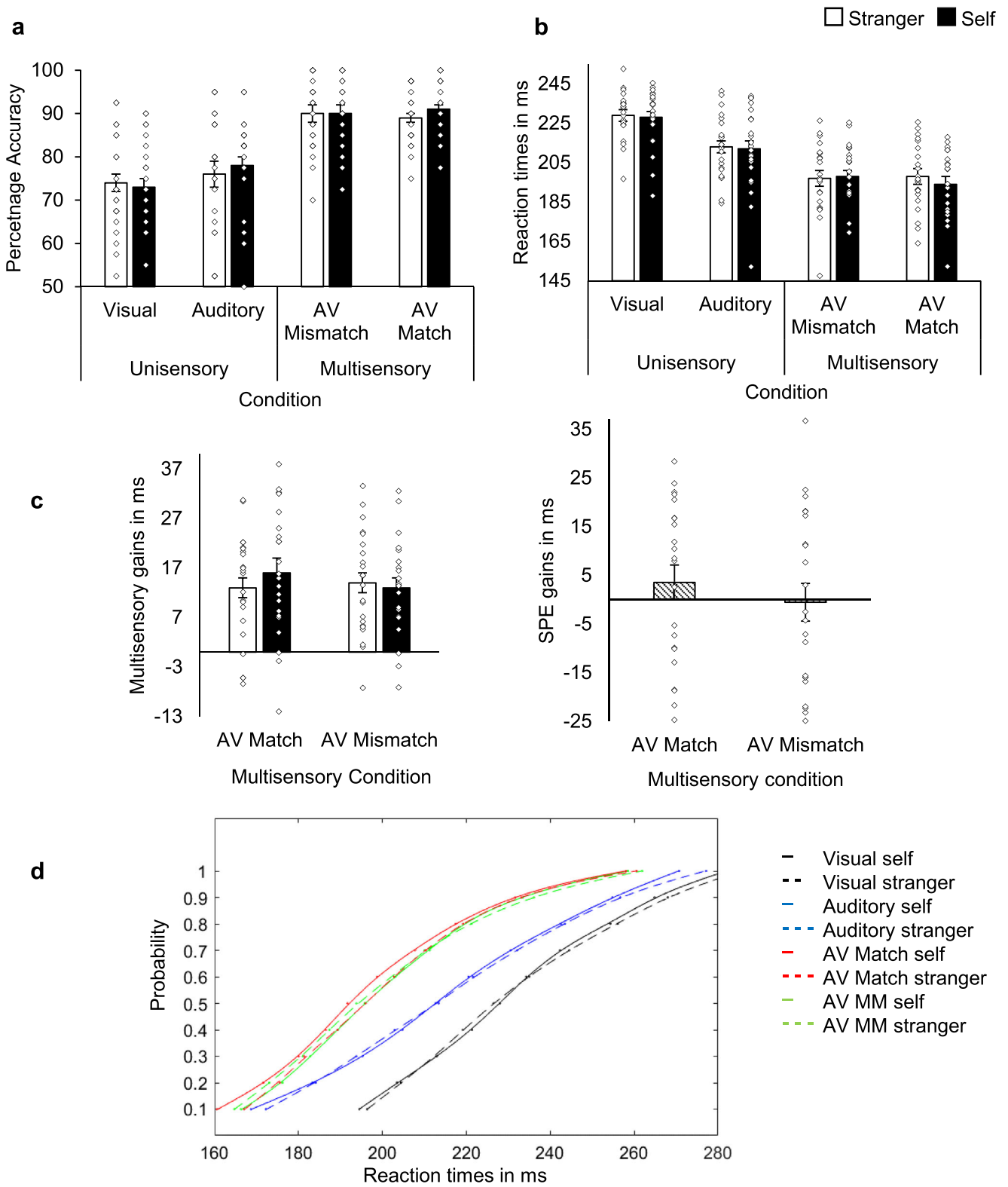
### RT

One condition was not normally distributed. Since the majority of the data were normally distributed, we chose not to transform the data. RTs were submitted to a 2 (Association: Self, Stranger)  $\times$  4 (Stimulus type: Visual, Auditory, AV Mismatch, AV Match) repeated-measures ANOVA. There was a significant main effect of Stimulus type,  $F(2.20, 46.23) = 115.95$ ,  $p < .001$ ,  $\eta_p^2 = .85$  (Greenhouse-Geisser correction). There were no other significant effects. Pairwise comparisons between stimulus types revealed a significant difference between all stimulus types ( $ps < .001$ ) except for between AV Mismatch and AV Match conditions ( $p = .14$ ). Responses in AV trials were faster than in unisensory trials, and responses in the Auditory trials were faster than in the Visual trials (see Fig. 5b).

### Multisensory facilitation in RT

Multisensory gain/cost values (faster of unisensory – multisensory) were submitted to a 2 (Association: self, stranger)  $\times$  2 (AV stimulus type: match, mismatch) repeated-measures ANOVA. There were no significant effects (Fig. 5c, left graph). SPE gains (see *Data analysis*) can also be seen in Fig. 5c (right graph). Multisensory percentage gains/costs were also calculated, the analysis re-run, and the findings were replicated.





**CDFs: RTs**

Three of the forty conditions (2 Association × 4 Stimulus type × 5 Probabilities) were not normally distributed, and there was one studentized residuals outlier. With the outlier excluded,

studentized residuals for two of the 40 conditions were not normally distributed. Since the majority of the data were normally distributed, we chose not to transform the data. With the outlier excluded (N = 21), RTs were submitted to a 2 (Association: self, stranger) × 4 (Stimulus type: A, V, AV

Mismatch, AV Match)  $\times$  5 (Probability: 1, 3, 5, 7, 9) repeated-measures ANOVA. There were main effects of Stimulus type,  $F(2.17, 43.40) = 117.55, p < .001, \eta_p^2 = .86$ , and Probability,  $F(1.34, 26.70) = 773.53, p < .001, \eta_p^2 = .98$ . There was a significant interaction between Stimulus type and Probability,  $F(5.48, 109.63) = 9.13, p < .001, \eta_p^2 = .31$ . (Greenhouse-Geisser correction used for all effects with more than two factors.) There were no other significant effects. Pairwise comparisons for Stimulus type revealed significant differences between the unisensory and multisensory conditions ( $ps < .001$ ), and the auditory ( $M = 215.17$  ms,  $SE = 2.28$ ) and visual conditions ( $M = 230.79$  ms,  $SE = 2.25$ ), ( $ps < .001$ ), but no significant difference between AV Mismatch ( $M = 199.37$  ms,  $SE = 2.42$ ) and AV Match ( $M = 197.84$  ms,  $SE = 2.69$ ) conditions,  $p = .11$  (see Fig. 5d). (The analysis was also conducted with the outlier included. The findings were replicated.)

The participant data exclusion criteria for the detection task analysis ( $N = 22$  sample) were based on participant performance in the detection task (irrespective of their performance in the matching task). This could potentially confound the transfer of the social associations because any participants with lower performance in the matching task may not have properly integrated the associations for access during the detection task. To check whether performance in the matching task could have impacted the influence of self-associations in the detection task, a supplementary analysis of RTs and multisensory percentage gains was conducted. This time the analysis excluded participant data that had been excluded from both the matching and the detection task analyses combined ( $N = 16$ ; see Appendix 4). The findings replicated the findings with the  $N = 22$  sample.

## Discussion

Using an audiovisual adaptation of Sui et al.'s (2012) matching task, the present study investigated the SPE – the magnitude of the difference in performance between motor responses to self-associated and stranger-associated stimuli. The SPE was modulated by whether the simultaneously presented label and object pairs were visual, auditory, or a combination of the two modalities. Specifically, the SPE in RT was significantly greater in responses to the visual shape and auditory label stimuli (V+AL) than to the auditory object and auditory label (A+AL) stimuli, and there was a significant interaction between association (self- vs. stranger-associated responses) and audiovisual stimulus type (A+VL vs. V+AL) for multisensory gains/costs in RT. Multisensory costs were descriptively greater for self- than stranger-associated A+VL stimuli, while for the V+AL stimuli, there were costs for stranger, and gains were uniquely observed for self. As such, the SPE interacted with the combination of the object and

label stimulus modalities. The SPE in RT was also diminished when stimuli were intermixed as compared with blocked by stimulus modality type. Furthermore, no significant self-advantage was found in simple detection task motor responses to the unisensory and multisensory stimuli in which the learnt self-associations were not relevant to the task at hand. Taken together, the present findings therefore indicate that the SPE can be modulated by both stimulus- and task-related parameters within the matching task, but the self-associations formed in the matching task do not automatically result in similar motor speed gains to unisensory and multisensory stimuli in fast ( $< 300$  ms) simple RT motor responses.

### SPE in the matching task

The present study findings are both consistent with, and also depart from, the findings reported by previous research. It was hypothesized that the SPE would be moderated by block type. Consistent with this prediction, and previous research (Golubickis & Macrae, 2021) block type moderated the SPE in RT. The SPE was reduced in intermixed trials (cf. the increased SPE in intermixed trials in the previous study in which visual stimuli were blocked or intermixed by the target shape identity). One possibility is that increasing fatigue may have reduced the SPE in intermixed trials since the intermixed condition was presented after the blocked condition. Conversely, the increased familiarity of the stimuli and with the S-R mappings gained in the blocked trials might be expected to increase the SPE in intermixed trials. An additional analysis was conducted, however, to examine the magnitude of the SPE across individual blocks within the blocked and intermixed conditions. The analysis confirmed that the SPE did not significantly decrease within either the blocked or the intermixed condition (see Appendix 6). Multisensory costs in self- as compared with stranger-associated responses were also differentially modulated by block type in the present study. Costs in RT were significantly reduced in blocked as compared with intermixed trials for self-associated responses, whereas costs in stranger-associated responses across block types did not differ. Taken together, these findings suggest that self- and stranger-associated responses were differentially modulated by block type.

It was also hypothesized, should findings be consistent with the pattern of results across previous studies examining the SPE using visual and auditory stimuli (Schäfer et al., 2016b; Stolte et al., 2021), that the SPE would be reduced with audiovisual as compared with visual stimuli, but that the latter may depend on the block type and the auditory stimulus intensity. Specifically, using intermixed trials and auditory stimuli presented at 75 dB, Stolte et al. (2021) documented that the SPE was diminished in those responses made to audiovisual stimuli (i.e., auditory tones with visual labels, and particularly diminished in the absence of labels with

audiovisual objects) as compared with visual-only stimuli with labels. Meanwhile, using blocked trials and auditory stimuli presented at 50 dB, Schäfer et al. (2016b) found that the SPE was of equivalent magnitude in responses to visual, auditory, and tactile stimuli paired with visual labels. In the present study, the SPE in RT was significantly attenuated with auditory as compared with visual objects (consistent with Stolte et al., 2021), but only when paired with auditory labels. (NB. Stolte et al. only paired their stimuli with visual labels.) Furthermore, there was an interaction between association (self- vs. stranger-associated responses) and audiovisual stimulus type (A+VL vs. V+AL) for multisensory gains/costs in RT. The multisensory (A+VL) stimuli were associated with ‘costs’ rather than gains over unisensory stimulation (in line with our expectations, and consistent with Stolte et al., 2021), with descriptively greater costs for self- than stranger-associated responses. However, while costs were also observed in participants’ responses to stranger-associated multisensory (V+AL) stimuli, multisensory gains were observed in responses to the self-associated multisensory (V+AL) stimuli. Thus, the present study provides the first evidence that the combination of the label and object modality (auditory vs. visual) can modulate the SPE.

Taken together, these patterns of results suggest that the SPE is influenced by both the task design (blocked vs. intermixed trials) and the stimulus parameters (the modality of the object and label stimuli). Contrary to predictions, however, auditory stimulus intensity did not moderate the SPE in the present study. Further analyses confirmed that there were no significant main nor interaction effects of auditory stimulus intensity on the SPE in the absolute RTs either – analysis reported in Appendix 5. It is important to note that a very narrow super-threshold dB range (50 dB and 70 dB) was used in order to replicate past studies (Schäfer et al., 2016b; Stolte et al., 2021). In future studies, comparing the SPE across a greater range of auditory stimulus intensities from near-threshold levels to upper limits may yield different results.

It was also hypothesized that a reduced SPE with audiovisual stimuli would lend support to Stolte et al.’s (2021) findings, whereas an equivalent SPE across stimulus modality types would lend support to the contention that the SPE is underpinned by a modulation of modality-general processes (in line with Schäfer et al., 2016b). In their study, Stolte et al. also demonstrated that the SPE and, in turn, visual dominance, was modulated by the relative physical hierarchy of the tones used (i.e., the multisensory SPE was akin to the visual SPE if the high tone was paired with the self). The authors suggested that it may, under certain conditions, be easier to form self-associations with visual representations of objects than audiovisual representations. Meanwhile, Schäfer et al. (2016b) have suggested that the SPE is likely to be a modality-general mechanism, since the SPE was of equivalent magnitude in responses to their visual, auditory, and tactile stimuli paired

with visual labels. As noted, the present study found that the magnitude of the SPE in the matching task was moderated across stimulus modality types, but was not consistently reduced with auditory stimuli. Thus, we did not find evidence to support the notion that the SPE is underpinned by modality-general processes (although this could not be ruled out), nor that it may be easier to form visual self-associations per se.

Schäfer et al.’s (2016b) findings and the results of the present study suggest that when the visual label is paired with an auditory or visual object, the self-advantage that arises is of equivalent magnitude. However, the use of auditory labels paints a rather different picture. In the present study, a significant difference between the SPE in responses to the A+AL and V+AL stimuli was documented. One interpretation for this asymmetry is that self-reference has the most marked advantage over stranger-reference in forming or accessing multisensory associations (specifically involving the visual shape and auditory label stimuli), and the weakest advantage in forming or accessing unisensory auditory associations. Whether self-associations involve sensory or modality-general representations or both is a question that is outside the scope of the present study (cf. Schäfer et al., 2020b). At the very least, when visual or auditory stimulus objects are paired with auditory labels, the present study suggests that the SPE is reduced with auditory stimuli.

Task-related factors, for example, that elicit active attention to visual over auditory stimuli, could also moderate the SPE across stimulus modality types. As noted, one possibility is that the alerting properties of auditory stimuli may elicit an override response in participants such that they actively focus their attention toward the weaker visual stimuli as a consequence. Fewer cognitive resources thus remain to attend to, and process, auditory stimuli (Posner et al., 1976). We also tend to rely more on visual than auditory information (Sinnett et al., 2007) and visual dominance is socially and culturally reinforced (Hutmacher, 2019). Parsing and extracting the semantic content of simultaneously-presented auditory object and label stimuli versus visual object and auditory label stimuli is more unusual. Cues to solve the former task may automatically be sought from the visual system, for example, modulating auditory stimulus processing, and thus the SPE. Determining the strategies underlying the processing of the object-label stimuli in the A+AL and V+AL conditions (e.g., using imaging methods such as functional magnetic resonance imaging (MRI) and electroencephalography (EEG)), and examining how self-relevance interacts with these processes, would likely provide further insight into the mechanisms underlying the SPE.

### SPE in the detection task

Task instructions and whether the meaning of the stimuli are relevant to the task at hand can also modulate both self-bias

and multisensory processes (Barutchu & Spence, 2021; Caughey et al., 2021; Dalmaso et al., 2019; Falbén et al., 2019; Macrae et al., 2017; Woźniak & Knoblich, 2021). Indeed, in contrast to the matching task, the predicted ‘classic’ multisensory gain was consistently observed in the detection paradigm, but no statistically significant self-advantage (i.e., SPE) was detected in either RTs or gain measures. Neural processes can rapidly change in anticipation of task-relevant stimuli in visual tasks (Corbetta et al., 2000; Nobre & van Ede, 2018; Stokes et al., 2009), and task instructions have been shown to alter multisensory gains and costs across consecutive tasks (Barutchu & Spence, 2021; Sinnett et al., 2008). Such top-down processes determined by task instruction may interact with the SPE and multisensory processes to modulate gains and costs to task-relevant and -irrelevant stimuli.

The absence of a significant SPE in the detection task, however, may also be related to other factors. The RT gains were much smaller than expected (on average in the order of 20 ms) as compared with past studies (e.g., Barutchu et al., 2009; Miller, 1982). Unlike in prior studies in which responses are typically limited to an 800- to 2,000-ms window, RTs were limited to 300 ms. This short response time limit was used to ensure that participants did not delay their responses (to avoid the intentional evaluation of the stimuli), and to keep participants alert to the task and responding as fast as they could, rather than simply instructing them to respond as fast as possible. This limited response window, and pressure to respond as rapidly as possible, may have reduced the RT gain measures.

Participants in Group 2 may also have been fatigued, with attention and motivation levels low during the detection task as it was always performed last to allow for optimal learning of stimulus associations. Multisensory processes and facilitation effects are partly dependent on attention (e.g., Barutchu et al., 2021; Talsma, 2015; Talsma et al., 2010; Zuanazzi & Noppeney, 2019). Our findings, however, are consistent with previous studies using other task paradigms that did not observe self-prioritization when the stimuli were not semantically evaluated (Caughey et al., 2021; Dalmaso et al., 2019; Falbén et al., 2019; Stein et al., 2016), or at least when the associations were not represented by the participant as task relevant (Woźniak & Knoblich, 2021). Furthermore, as noted earlier, additional analyses (see Appendix 6) suggested that the SPE did not significantly decrease across individual blocks within either the blocked condition or the intermixed condition in the preceding Matching task.

One further consideration is that in the present study geometric shapes were used as visual stimuli, while the sound stimuli were musical instrument bursts, in order to keep the stimuli consistent with previous research (Schäfer et al., 2016b; Stolte et al., 2021). Alternatively, drawings of musical instruments could be used as visual stimuli (i.e., a different exemplar of the same category; Schäfer et al., 2015). Using

different categories across visual and auditory stimuli may have amplified the working memory load in intermixed trials and further reduced the SPE in the present study. Notably, however, if increased working memory load reduces the SPE, then a reduced SPE would still be expected in intermixed trials even with conceptually more similar stimuli. Furthermore, the short 150-ms bursts were not typical of the characteristic sounds made by those instruments, and although recognisable as belonging to their respective categories (when narrowed down to two possibilities), the sound bursts were not typical of the instruments or familiar. This would be more likely to facilitate the forging of self/other associations with sensory rather than conceptual features of the sounds. Conversely, using drawings of the instruments and sounds of the instruments would encourage the use of ‘conceptual’ associations as this would be the optimal strategy to maximise responding across modality types in intermixed trials. Future studies could examine whether matching the semantic content of the stimuli (other than the person associations) across stimulus types modulates the SPE in intermixed trials.

## Conclusion

The present study found that the SPE in an audiovisual adaptation of Sui et al.’s (2012) matching task was diminished in intermixed as compared with trials blocked by stimulus modality type, and interacted with label and object modality (auditory or visual). The standard-sized self-advantage did not arise in simple detection motor responses to the unisensory and multisensory stimuli. Taken together, these findings indicate that the SPE is modulated by both stimulus- and task-related parameters within the matching task, and that the self-associations formed in the matching task do not automatically result in similar motor speed gains to unisensory and multisensory stimuli in a simple detection task.

## Appendix 1

**Absolute RTs and sensitivity index scores (D-prime;  $d'$ ) (Group 1).** There was one studentized residuals outlier (outside  $\pm 3.0$  in absolute value in one or more of the conditions) in RT. With the outlier removed ( $N = 27$ ), the matching-trial RT data were submitted to a 2 (Block type: blocked, intermixed)  $\times$  4 (Stimulus type: V-shape+VLabel, A+AL, A+VL, V+AL)  $\times$  2 (Association: Self, Stranger) repeated-measures ANOVA. The analysis revealed significant main effects of Block type,  $F(1, 26) = 191.85, p < .001, \eta_p^2 = .88$ , Stimulus type,  $F(2.186, 56.849) = 22.61, p < .001, \eta_p^2 = .47$ , and Association,  $F(1, 26) = 43.07, p < .001, \eta_p^2 = .62$ . RTs were shorter in the Blocked ( $M = 705$  ms,  $SE = 9.24$ ) as



compared with the Intermixed trials ( $M = 781$  ms,  $SE = 9.49$ ). There were significant differences in RTs between all Stimulus types ( $ps < .004$ ), except between V+VL and V+AL ( $p = .10$ ). RTs in self ( $M = 715$  ms,  $SE = 9.58$ ) as compared with stranger-associated trials were shorter ( $M = 771$  ms,  $SE = 10.26$ ), indicating the presence of a SPE.

There was a significant interaction between Block type and Stimulus type,  $F(3, 78) = 8.55$ ,  $p < .001$ ,  $\eta_p^2 = .25$ . There were no other significant interaction effects. Probing the interaction between Block type and Stimulus type revealed a significant difference between blocked and intermixed trial responses across all stimulus types ( $ps < .001$ ). RTs were shorter in the blocked as compared with the intermixed stimulus type trials. In the blocked trials, only the difference in RTs between A+AL and V+AL was significant ( $p < .001$ ). In contrast, in intermixed trials, there was a significant difference between each of the stimulus types ( $ps < .001$ ) except for between V+VL and V+AL ( $p = .34$ ). (The analysis was conducted again with the outlier included and the findings were replicated.)

The  $d'$  data were normally distributed except for one condition (Blocked V+AL Stranger), therefore we chose not to transform the data. Sensitivity index scores were submitted to a 2 (Block type: blocked, intermixed)  $\times$  4 (Stimulus type: V-shape+VLabel, A+AL, A+VL, V+AL)  $\times$  2 (Association: Self, Stranger) repeated-measures ANOVA. There were significant main effects of Block type,  $F(1, 27) = 29.26$ ,  $p < .001$ ,  $\eta_p^2 = .52$ , Stimulus type,  $F(2.183, 58.943) = 15.67$ ,  $p < .001$ ,  $\eta_p^2 = .37$  (Greenhouse-Geisser correction), and Association,  $F(1, 27) = 12.89$ ,  $p = .001$ ,  $\eta_p^2 = .32$ .  $d'$  scores were higher for responses to Blocked ( $M = 2.56$ ,  $SE = .09$ ) as compared with Intermixed ( $M = 1.94$ ,  $SE = .14$ ) trials. There were significant differences between all Stimulus types ( $ps < .01$  for all) except for between V+VL and A+VL ( $p = .76$ ).  $d'$  scores were higher for self- ( $M = 2.38$ ,  $SE = .12$ ) as compared with stranger-associated ( $M = 2.11$ ,  $SE = .11$ ) responses, indicating the presence of a SPE.

**Absolute RTs and sensitivity index scores (D-prime;  $d'$ ) (Group 2).** There was one studentized residuals outlier (outside  $\mp 3.0$  in absolute value) in RT. With the outlier removed ( $N = 21$ ), studentized residuals were normally distributed except for in two of the conditions. Since the majority of the data were normally distributed, we chose not to transform the data. The matching trial RT data ( $N = 21$ ) were submitted to a 2 (Block type: blocked, intermixed)  $\times$  4 (Stimulus type: V+VL, A+AL, A+VL, V+AL)  $\times$  2 (Association: Self, Stranger) repeated measures ANOVA. There were significant main effects of Block type,  $F(1, 20) = 53.03$ ,  $p < .001$ ,  $\eta_p^2 = .73$ , Stimulus type,  $F(3, 60) = 19.83$ ,  $p < .001$ ,  $\eta_p^2 = .50$ , and Association,  $F(1, 20) = 6.05$ ,  $p = .02$ ,  $\eta_p^2 = .23$ . In contrast to the results for Group 1, there were also significant simple two-way interactions between Block type and Association,  $F(1, 20) = 9.72$ ,  $p = .01$ ,  $\eta_p^2 = .33$ , and between Stimulus type and Association,  $F(3, 60) = 4.30$ ,  $p = .008$ ,  $\eta_p^2 = .18$ , and, as

for Group 1, between Block type and Stimulus type,  $F(3, 60) = 3.64$ ,  $p = .02$ ,  $\eta_p^2 = .15$ . Probing the interaction between Stimulus type and Association revealed a significant difference between self- and stranger-associated responses for the V+VL ( $p = .005$ ,  $\eta_p^2 = .34$ ) and V+AL ( $p = .008$ ,  $\eta_p^2 = .30$ ) stimulus types, but not for the A+VL ( $p = .07$ ,  $\eta_p^2 = .16$ ) and A+AL ( $p = .39$ ,  $\eta_p^2 = .04$ ) stimulus types. Probing the interaction between Block type and Association revealed a significant difference between self- and stranger-associated responses in the Blocked trials,  $F(1, 21) = 9.20$ ,  $p = .007$ ,  $\eta_p^2 = .32$ , but not in the intermixed trials,  $F(1, 21) = 2.64$ ,  $p = .12$ ,  $\eta_p^2 = .12$ . Once again, these findings were replicated with the outlier included.

Sensitivity index scores were submitted to a 2 (Block type: blocked, intermixed)  $\times$  4 (Stimulus type: V+VL, A+AL, A+VL, V+AL)  $\times$  2 (Association: Self, Stranger) repeated-measures ANOVA. The analysis revealed significant main effects of Block type,  $F(1, 21) = 40.61$ ,  $p < .001$ ,  $\eta_p^2 = .66$ , and Stimulus type,  $F(3, 63) = 14.28$ ,  $p < .001$ ,  $\eta_p^2 = .41$ , but, in contrast to Group 1, not of Association,  $F(1, 21) = 1.90$ ,  $p = .18$ ,  $\eta_p^2 = .08$ . Self- ( $M = 2.31$ ,  $SE = .16$ ). Self- as compared with stranger- ( $M = 2.15$ ,  $SE = .20$ ) associated  $d'$  scores were not significantly different. In contrast to Group 1, there was a significant interaction between Block type and Stimulus type,  $F(3, 63) = 4.95$ ,  $p = .004$ ,  $\eta_p^2 = .19$ .

## Appendix 2

**Matching tasks (Groups 1 and 2).** An independent  $t$ -test was conducted on overall RT across Groups 1 and 2 revealing no significant difference between the 50-dB auditory group ( $M = 742$  ms,  $SD = 46$  ms) and the 70-dB stimulus intensity group ( $M = 730$  ms,  $SD = 54$  ms),  $t(48) = 0.82$ ,  $p = .42$ . Similarly, there was no significant difference between overall  $d'$  scores between the 50-dB group ( $M = 2.25$ ,  $SD = 0.56$ ) and the 70-dB group ( $M = 2.23$ ,  $SD = 0.79$ ),  $t(36.31) = 0.07$ ,  $p = .94$ .

A 2 (Auditory stimulus intensity: 50 dB, 70 dB)  $\times$  2 (Block type: Blocked, Intermixed)  $\times$  4 (Stimulus type: V+VL, A+AL, A+VL, V+AL) ANOVA on RTs across Groups 1 and 2 with the Association (self, stranger) condition collapsed revealed no effect of auditory stimulus intensity. In particular, there was no interaction between Block type and Auditory stimulus intensity ( $p = .31$ ), nor between Stimulus type and Auditory stimulus intensity ( $p = .57$ ), nor was there a three-way interaction ( $p = 1.00$ ).

A 2 (Auditory stimulus intensity: 50 dB, 70 dB)  $\times$  2 (Block type: Blocked, Intermixed)  $\times$  4 (Stimulus type: V+VL, A+AL, A+VL, V+AL) ANOVA on  $d'$  scores across Groups 1 and 2 with the Association (self, stranger) condition collapsed revealed no effect of auditory stimulus intensity: There was no interaction between Block type and Auditory stimulus intensity,  $p = .90$ , nor between Stimulus type and Auditory stimulus intensity ( $p = .72$ ), nor was there a three-way interaction ( $p = .44$ ).

## Appendix 3

The following analyses replicate those presented in the main text, but with the outliers included (rather than excluded) in each case. The findings reported in the main text were all replicated.

### Self-bias in RTs with outliers included

F-statistics are presented in Table A1. There were two outliers (studentized residuals outside  $\mp 3.0$  in absolute value) identified in the self-bias index scores (there was one outlier which when removed revealed a second outlier). The majority of the data were normally distributed (three conditions not normally distributed), so we chose not to transform the data. The assumption of homogeneity of variances was violated for one condition, as assessed by Levene's test for equality of variances. Given that the sample sizes of the two groups were roughly equal (ratio 1.29) and ANOVA is generally robust to violations of this assumption if group sizes are roughly equal, ANOVA was carried out on the data. With the outliers included, a 2 (Auditory stimulus intensity: 50 dB, 70 dB)  $\times$  2 (Block type: Blocked, Intermixed)  $\times$  4 (Stimulus type: V+VL, A+AL, A+VL, V+AL) mixed ANOVA was conducted on the RT self-bias index scores. The analysis revealed main effects of Stimulus type, and Block type. Overall, the magnitude of self-bias in the Blocked trials ( $M = .04$ ,  $SE = .007$ ) was greater than in the Intermixed trials ( $M = .02$ ,  $SE = .005$ ). For Stimulus type, the magnitude of the normalised self-bias in the RT data was significantly greater ( $p = .002$ ) in responses to V+AL ( $M = .04$ ,  $SE = .007$ ) as compared with A+AL ( $M = .02$ ,  $SE = .006$ ). There were no significant differences between any other pair of Stimulus types ( $ps > .02$ ). There were no other significant effects. These findings replicated those with the outliers excluded – see main text for details.

### Self-bias in sensitivity index scores (D-prime; $d'$ ) with outliers included

F-statistics are presented in Table A1. There were two studentized residuals outliers (studentized residuals outside  $\mp 3.0$  in absolute value) identified in the self-bias index scores (there was one outlier which when removed revealed a second outlier). The assumption of homogeneity of variances was violated for two conditions. As with the analysis of self-bias in the RT data, given that the group sizes were roughly equal, an ANOVA was carried out on the data ( $N = 50$ ). With the outliers included, a 2 (Auditory stimulus intensity: 50 dB, 70 dB)  $\times$  2 (Block type: Blocked, Intermixed)  $\times$  4 (Stimulus type: V+VL, A+AL, A+VL, V+AL) mixed ANOVA was conducted on self-bias index scores for  $d'$  (see Table 1 – main text). There were no significant effects. The SPE in sensitivity between self- and stranger-associated responses did not differ significantly across the conditions. These findings replicated those with the outliers excluded – see main text.

### Multisensory percentage gains/costs in RT with outliers included

F-statistics and pairwise comparisons are presented in Table A2. There were two studentized residuals outliers whose removal produced a third outlier. With the outliers included, the data were submitted to a 2 (Auditory stimulus intensity: 50 dB, 70 dB)  $\times$  2 (Block type: blocked, intermixed)  $\times$  2 (AV Stimulus type: A+VL, V+AL)  $\times$  2 (Association: self, stranger) mixed ANOVA. There was a significant main effect of Block type, and AV Stimulus type. Negative gain values indicated that there were costs for responses to multisensory relative to unisensory stimuli. Multisensory costs were greater in participants' responses to intermixed trials ( $M = -4.99$  ms,  $SE = 0.75$ ) as compared with blocked trials ( $M = -1.92$  ms,  $SE = 0.98$ ), and greater in responses to the A+VL stimulus type ( $M = -6.72$ ,  $SE = 0.94$ ) as compared with responses to the V+AL stimulus type ( $M = -0.20$  ms,  $SE = 0.78$ ). There was a significant interaction between Block type and Association, and between AV Stimulus type and Association. None of the other effects was significant.

Probing the interaction between AV stimulus type and Association revealed a significant difference between stranger-associated responses to the A+VL ( $M = -5.88$  ms,  $SE = 1.14$ ) and V+AL ( $M = -1.19$  ms,  $SE = 0.94$ ) stimulus types, and a significant difference between self-associated responses to the A+VL ( $M = -7.55$  ms,  $SE = 1.10$ ) and V+AL stimulus types ( $M = 0.79$  ms,  $SE = 0.92$ ). There was no significant difference between self- and stranger-associated responses to the V+AL stimulus type, or the A+VL stimulus type.

Probing the Block type and Association interaction revealed no significant difference in multisensory costs for self-associated responses ( $M = -5.86$  ms,  $SE = 0.78$ ) than for with stranger-associated responses ( $M = -4.13$  ms,  $SE = 0.96$ ) in intermixed trials. There was no significant difference between self-associated responses ( $M = -0.90$  ms,  $SE = 1.17$ ) as compared with stranger-associated responses ( $M = -2.94$  ms,  $SE = 1.11$ ) in the blocked trials, but it is worth noting that costs were descriptively greater in stranger-associated as compared with self-associated responses. In addition, stranger-associated responses in blocked trials as compared with intermixed trials were not significantly different. In contrast, multisensory costs in self-associated responses in blocked trials as compared with intermixed trials were significantly different. Multisensory costs in self- as compared with stranger-associated responses were differentially modulated by block type. These findings replicated those with the outliers excluded – see main text, except that there was no significant difference between multisensory costs for self- versus stranger-associated responses in intermixed trials.

### Multisensory gains/costs in sensitivity index scores (D-prime; $d'$ ) with outliers included

F-statistics and pairwise comparisons are presented in Table A2. There were four studentized residuals outliers.

With the outliers included, the data were submitted to a 2 (Auditory stimulus intensity: 50 dB, 70 dB)  $\times$  2 (Block type: blocked, intermixed)  $\times$  2 (AV stimulus type: A+VL, V+AL)  $\times$  2 (Association: self, stranger) mixed ANOVA. The analysis revealed a significant main effect of AV Stimulus type. Negative gain values indicated that there were costs for responses to multisensory relative to unisensory stimuli. There was some multisensory facilitation in responses to the V+AL stimuli ( $M = 0.06$ ,  $SE = 0.08$ ), and costs in responses to A+VL stimuli ( $M = -0.38$ ,  $SE = 0.08$ ). There was a significant interaction between Block type and AV stimulus type. There were no other significant effects. The interaction between Block type and AV stimulus type was probed revealing that multisensory costs in responses to the A+VL stimulus type ( $M = -0.25$ ,  $SE = 0.13$ ) as compared with the V+AL stimulus type ( $M = 0.01$ ,  $SE = 0.13$ ) in blocked trials were significantly different. There were multisensory gains in responses to the V+AL stimulus type ( $M = 0.10$ ,  $SE = 0.09$ ) as compared with multisensory costs in responses to the A+VL stimulus type ( $M = -0.51$ ,  $SE = 0.11$ ) in the intermixed trials. (These findings replicated those with the outliers excluded – see main text, except that there was a significant difference between multisensory costs in responses to the A+VL and V+AL stimulus types with the outliers included.)

## Appendix 4

**Multisensory simple detection task (Group 2).** The detection task sample included data from those participants whose data were excluded from the preceding matching task because they achieved below-threshold accuracy. This could potentially confound the transfer of the social associations because these participants may not have properly assimilated them. To check whether the low accuracy performance in the matching task could have impacted the influence of self-associations in the detection task, a further analysis was also conducted.

### Participants

Twenty-six participants completed the detection task. The data from four participants were excluded for having  $>2.5$  SDs below the group percentage accuracy mean for any one condition, or for achieving  $<50\%$  accuracy in any one condition in the detection task. In addition, the data from the seven participants that were excluded from analyses of the matching task data for low performance accuracy or constituting outliers (see main text) were also excluded. The data from 16 participants were used in the analysis.

### RTs

The RT data were submitted to a 2 (Association: Self, Stranger)  $\times$  4 (Stimulus type: Visual, Auditory, AV Mismatch, AV Match) repeated-measures ANOVA. There was a significant main effect of Stimulus type,  $F(2.05, 30.81) = 105.01$ ,  $p < .001$ ,  $\eta_p^2 = .88$  (Greenhouse-Geisser

correction). There was no main effect of Association,  $F(1, 15) = .15$ ,  $p = .70$ ,  $\eta_p^2 = .01$ . Pairwise comparisons between stimulus types revealed a significant difference between the unisensory and multisensory stimuli, and between the visual and auditory stimuli ( $ps < .001$ ). Note that the findings using the  $N = 16$  sample thus replicated the findings using the  $N = 22$  sample.

### Multisensory percentage gains/costs

Multisensory percentage gain/cost values ( $[(\text{faster of unisensory} - \text{multisensory}) / \text{faster of unisensory}] * 100$ ); see *Data analysis*) were submitted to a 2 (Association: self, stranger)  $\times$  2 (AV stimulus type: match, mismatch) repeated-measures ANOVA. There were no significant effects. The findings using the  $N = 16$  sample thus replicated the findings using the  $N = 22$  sample.

## Appendix 5

Auditory stimulus intensity did not moderate the SPE index scores. To confirm that there were no main nor interaction effects of auditory stimulus intensity on the SPE in the raw RTs, an additional ANOVA was carried out. The majority of the RT data were normally-distributed (two conditions not normally distributed), so we chose not to transform the data. There was one studentized residuals outlier. The assumption of homogeneity of variances was violated for one condition, as assessed by Levene's test for equality of variances. As with the analysis of self-bias in the RT data, given that the group sizes were roughly equal, an ANOVA was carried out on the data. With the outlier included ( $N = 48$ ), RTs were submitted to a 2 (between-groups Intensity: 50 dB, 70 dB)  $\times$  2 (within-groups Block type: blocked, intermixed)  $\times$  4 (within groups Stimulus type: V+VL, A+AL, A+VL, V+AL)  $\times$  2 (Within-groups Association: self, stranger) mixed factorial ANOVA. The analysis revealed main effects of Stimulus type,  $F(2.46, 113.10) = 40.81$ ,  $p < .001$ ,  $\eta_p^2 = .47$ , Block type,  $F(1, 46) = 187.56$ ,  $p < .001$ ,  $\eta_p^2 = .80$ , and Association,  $F(1, 46) = 55.09$ ,  $p < .001$ ,  $\eta_p^2 = .55$ . There was a significant interaction between Block type and Stimulus type,  $F(3, 138) = 10.10$ ,  $p < .001$ ,  $\eta_p^2 = .18$ , Block type and Association,  $F(1, 46) = 15.60$ ,  $p < .001$ ,  $\eta_p^2 = .25$ , and Stimulus type and Association,  $F(2.20, 100.97) = 4.34$ ,  $p = .01$ ,  $\eta_p^2 = .09$ . There was no interaction between Association and Auditory stimulus intensity,  $p = .40$ , between Block type, Association and Auditory stimulus intensity,  $p = .42$ , between Stimulus type, Association, and Auditory stimulus intensity,  $p = .53$ , or between Block type, Stimulus type, Association, and Auditory stimulus intensity,  $p = .53$ . There were no other significant effects. These findings confirmed that there are no main nor interaction effects of auditory stimulus intensity on the SPE in the raw RTs.



## Appendix 6

In the present study, participants completed blocks of trials blocked by stimulus modality type prior to completing the blocks of trials with all the stimulus modality types intermixed. Familiarity with both the matching task and the 16 stimulus-response mappings ( $2 \times \text{Association} \times 2 \text{ Match-type} \times 4 \text{ stimulus types}$ ) gained during the blocked trials enabled participants to complete the blocks of intermixed trials which were otherwise too difficult. Notably, the increased familiarity of the stimuli and the S-R mappings gained in the blocked trials might be expected to increase the SPE in intermixed trials. However, additional analyses were conducted to test whether the magnitude of the SPE significantly decreased across individual blocks within each of the Blocked and Intermixed conditions of the Matching task. Given that Stimulus modality type was balanced across individual blocks, and our previous analyses indicated that Auditory stimulus intensity had no significant effect, the Stimulus modality type (V+VL, A+AL, A+VL, V+AL) and Auditory stimulus intensity (50 dB, 70dB) conditions were collapsed. Self-bias index scores in RT were then submitted to a one-way repeated-measures ANOVA (Block number: 1, 2, 3, 4) to compare the SPE across the individual blocks of trials blocked by stimulus modality type. There was one studentized residuals outlier (studentized residuals outside  $\mp 3.0$  in absolute value), and one condition was not normally distributed. With the outlier included ( $N = 48$ ), the analysis revealed no significant main effect of Block,  $F(3, 141) = 2.40, p = .07, \eta_p^2 = .05$ , indicating that the magnitude of the SPE did not decrease significantly across blocks. With the outlier excluded from the analysis ( $N = 47$ ), this finding was replicated,  $F(3, 138) = 2.38, p = .07, \eta_p^2 = .05$ .

Self-bias index scores in RT were then submitted to a one-way repeated-measures ANOVA (Block number: 1, 2, 3, 4, 5) to compare the SPE across the individual blocks of intermixed trials ( $N = 48$ ). There was one studentized residuals outlier, and one condition was not normally distributed. With the outlier included ( $N = 48$ ), the analysis revealed no significant main effect of Block,  $F(1.21, 56.94) = 0.54, p = .70, \eta_p^2 = .01$  (Greenhouse-Geisser correction), indicating that the magnitude of the SPE did not decrease significantly across blocks. With the outlier excluded from the analysis ( $N = 47$ ), all conditions were normally distributed, and this finding was replicated,  $F(3.33, 153.10) = 1.40, p = .24, \eta_p^2 = .03$  (Greenhouse-Geisser correction).

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.3758/s13414-022-02498-z>.

**Funding** This study was part-funded by a bursary from the Eurofins Foundation awarded to the first author.

**Open practices statement** The datasets generated and analysed during the current study are available from the corresponding author upon reasonable request. The experiments were not formally preregistered.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Alexopoulos, T., Muller, D., Ric, F., & Marendaz, C. (2012). I, me, mine: Automatic attentional capture by self-related stimuli. *European Journal of Social Psychology, 42*(6), 770–779. <https://doi.org/10.1002/ejsp.1882>
- Arana, S., Marquand, A., Hultén, A., Hagoort, P., & Schoffelen, J.-M. (2020). Sensory modality-independent activation of the brain network for language. *Journal of Neuroscience, 40*(14), 2914–2924. <https://doi.org/10.1523/JNEUROSCI.2271-19.2020>
- Barutchu, A., & Spence, C. (2021). Top-down task-specific determinants of multisensory motor reaction time enhancements and sensory switch costs. *Experimental Brain Research, 239*(3), 1021–1034. <https://doi.org/10.1007/s00221-020-06014-3>
- Barutchu, A., Crewther, D. P., & Crewther, S. G. (2009). The race that precedes coactivation: Development of multisensory facilitation in children. *Developmental Science, 12*(3), 464–473. <https://doi.org/10.1111/j.1467-7687.2008.00782.x>
- Barutchu, A., Danaher, J., Crewther, S. G., Innes-Brown, H., Shivdasani, M. N., & Paolini, A. G. (2010). Audiovisual integration in noise by children and adults. *Journal of Experimental Child Psychology, 105*(1), 38–50. <https://doi.org/10.1016/j.jecp.2009.08.005>
- Barutchu, A., Spence, C., & Humphreys, G. W. (2018). Multisensory enhancement elicited by unconscious visual stimuli. *Experimental Brain Research, 236*(2), 409–417. <https://doi.org/10.1007/s00221-017-5140-z>
- Barutchu, A., Fifer, J. M., Shivdasani, M. N., Crewther, S. G., & Paolini, A. G. (2020). The interplay between multisensory associative learning and IQ in children. *Child Development, 91*(2), 620–637. <https://doi.org/10.1111/cdev.13210>
- Belin, P., Fillion-Bilodeau, S., & Gosselin, F. (2008). The Montreal Affective Voices: A validated set of nonverbal affect bursts for research on auditory affective processing. *Behavior Research Methods, 40*(2), 531–539. <https://doi.org/10.3758/BRM.40.2.531>
- Brédart, S., Delchambre, M., & Laureys, S. (2006). One's own face is hard to ignore. *Quarterly Journal of Experimental Psychology, 59*(1), 46–52. <https://doi.org/10.1080/17470210500343678>
- Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concrete ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods, 46*(3), 904–911. <https://doi.org/10.3758/s13428-013-0403-5>
- Campbell, J. I. D., & Thompson, V. A. (2012). MorePower 6.0 for ANOVA with relational confidence intervals and Bayesian analysis. *Behavior Research Methods, 44*(4), 1255–1265. <https://doi.org/10.3758/s13428-012-0186-0>



- Caughy, S., Falbén, J. K., Tsamadi, D., Persson, L. M., Golubickis, M., & Macrae, N. C. (2021). Self-prioritization during stimulus processing is not obligatory. *Psychological Research*, 85(2), 503–508. <https://doi.org/10.1007/s00426-019-01283-2>
- Chen, Y.-C., & Spence, C. (2018). Audiovisual semantic interactions between linguistic and non-linguistic stimuli: The time-courses and categorical specificity. *Journal of Experimental Psychology: Human Perception and Performance*, 44(10), 1488–1507. <https://doi.org/10.1037/xhp0000545>
- Chiarella, S. G., Makwana, M., Simone, L., Hartkamp, M., Calabrese, L., Raffone, A., & Srinivasan, N. (2020). Mindfulness meditation weakens attachment to self: Evidence from a self vs. other binding task. *Mindfulness*, 11(10), 2411–2422. <https://doi.org/10.1007/s12671-020-01457-9>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). L. Erlbaum Associates.
- Constable, M. D., Kritikos, A., & Bayliss, A. P. (2011). Grasping the concept of personal property. *Cognition*, 119(3), 430–437. <https://doi.org/10.1016/j.cognition.2011.02.007>
- Constable, M. D., Kritikos, A., Lipp, O. V., & Bayliss, A. P. (2014). Object ownership and action: The influence of social context and choice on the physical manipulation of personal property. *Experimental Brain Research*, 232(12), 3749–3761. <https://doi.org/10.1007/s00221-014-4063-1>
- Constable, M. D., Welsh, T. N., Huffman, G., & Pratt, J. (2019). I before U: Temporal order judgements reveal bias for self-owned objects. *Quarterly Journal of Experimental Psychology*, 59, 46–52. <https://doi.org/10.1177/1747021818762010>
- Constable, M. D., Becker, M. L., Oh, Y.-I., & Knoblich, G. (2021). Affective compatibility with the self modulates the self-prioritisation effect. *Cognition and Emotion*, 35(2), 291–304. <https://doi.org/10.1080/02699931.2020.1839383>
- Conway, A. R. A., Cowan, N., & Bunting, M. F. (2001). The cocktail party phenomenon revisited: The importance of working memory capacity. *Psychonomic Bulletin & Review*, 8(2), 331–335. <https://doi.org/10.3758/BF03196169>
- Corbetta, M., Kincade, J. M., Ollinger, J. M., McAvoy, M. P., & Shulman, G. L. (2000). Voluntary orienting is dissociated from target detection in human posterior parietal cortex. *Nature Neuroscience*, 3(3), 292–297. <https://doi.org/10.1038/73009>
- Cunningham, S. J., & Turk, D. J. (2017). Editorial: A review of self-processing biases in cognition. *Quarterly Journal of Experimental Psychology*, 70(6), 987–995. <https://doi.org/10.1080/17470218.2016.1276609>
- Dalmaso, M., Castelli, L., & Galfano, G. (2019). Self-related shapes can hold the eyes. *Quarterly Journal of Experimental Psychology*, 72(9), 2249–2260. <https://doi.org/10.1177/1747021819839668>
- Desebrock, C., & Spence, C. (2021). The Self-Prioritization Effect: Self-referential processing in movement highlights modulation at multiple stages. *Attention, Perception, & Psychophysics*. <https://doi.org/10.3758/s13414-021-02295-0>
- Desebrock, C., Sui, J., & Spence, C. (2018). Self-reference in action: Arm-movement responses are enhanced in perceptual matching. *Acta Psychologica*, 190, 258–266. <https://doi.org/10.1016/j.actpsy.2018.08.009>
- Enock, F. E., Hewstone, M. R. C., Lockwood, P. L., & Sui, J. (2020). Overlap in processing advantages for minimal ingroups and the self. *Scientific Reports*, 10(1), 18933. <https://doi.org/10.1038/s41598-020-76001-9>
- Falbén, J. K., Golubickis, M., Balseryte, R., Persson, L. M., Tsamadi, D., Caughy, S., & Macrae, N. C. (2019). How prioritized is self-prioritization during stimulus processing? *Visual Cognition*, 27(1), 46–51. <https://doi.org/10.1080/13506285.2019.1583708>
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G\*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41(4), 1149–1160. <https://doi.org/10.3758/BRM.41.4.1149>
- Golubickis, M., & Macrae, C. N. (2021). Judging me and you: Task design modulates self-prioritization. *Acta Psychologica*, 218, 103350. <https://doi.org/10.1016/j.actpsy.2021.103350>
- Golubickis, M., Falbén, J. K., Sahraie, A., Visokomogilski, A., Cunningham, W. A., Sui, J., & Macrae, C. N. (2017). Self-prioritization and perceptual matching: The effects of temporal construal. *Memory & Cognition*, 45(7), 1223–1239. <https://doi.org/10.3758/s13421-017-0722-3>
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. John Wiley.
- Hecht, D., Reiner, M., & Kami, A. (2008). Multisensory enhancement: Gains in choice and in simple response times. *Experimental Brain Research*, 189(2), 133–143. <https://doi.org/10.1007/s00221-008-1410-0>
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2), 65–70.
- Houghton, G., & Grange, J. A. (2011). CDF-XL: Computing cumulative distribution functions of reaction time data in Excel. *Behavior Research Methods*, 43(4), 1023–1032. <https://doi.org/10.3758/s13428-011-0119-3>
- Hu, C.-P., Lan, Y., Macrae, C. N., & Sui, J. (2020). Good me bad me: Prioritization of the good-self during perceptual decision-making. *Collabra. Psychology*, 6(1), 20. <https://doi.org/10.1525/collabra.301>
- Humphreys, G. W., & Sui, J. (2015). The salient self: Social saliency effects based on self-bias. *Journal of Cognitive Psychology*, 27(2), 129–140. <https://doi.org/10.1080/20445911.2014.996156>
- Humphreys, G. W., & Sui, J. (2016). Attentional control and the self: The self-attention network. *Cognitive Neuroscience*, 7(1–4), 5–17. <https://doi.org/10.1080/17588928.2015.1044427>
- Hutmacher, F. (2019). Why is there so much more research on vision than on any other sensory modality? *Frontiers in Psychology*, 10, 1–12. <https://doi.org/10.3389/fpsyg.2019.02246>
- Kreutzfeldt, M., Stephan, D. N., Sturm, W., Willmes, K., & Koch, I. (2015). The role of crossmodal competition and dimensional overlap in crossmodal attention switching. *Acta Psychologica*, 155, 67–76. <https://doi.org/10.1016/j.actpsy.2014.12.006>
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, 4, 863. <https://doi.org/10.3389/fpsyg.2013.00863>
- Lukas, S., Andrea, M., & Koch, P. I. (2010). Switching attention between modalities: further evidence for visual dominance. *Psychological Research*, 74(3), 255–267. <https://doi.org/10.1007/s00426-009-0246-y>
- Ma, W. J., Zhou, X., Ross, L. A., Foxe, J. J., & Parra, L. C. (2009). Lip-reading aids word recognition most in moderate noise: A Bayesian explanation using high-dimensional feature space. *PLoS ONE*, 4(3), e4638. <https://doi.org/10.1371/journal.pone.0004638>
- Macrae, C. N., Visokomogilski, A., Golubickis, M., Cunningham, W. A., & Sahraie, A. (2017). Self-relevance prioritizes access to visual awareness. *Journal of Experimental Psychology: Human Perception and Performance*, 43(3), 438–443. <https://doi.org/10.1037/xhp0000361>
- Miller, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology*, 14(2), 247–279. [https://doi.org/10.1016/0010-0285\(82\)90010-X](https://doi.org/10.1016/0010-0285(82)90010-X)
- Miller, J., Franz, V., & Ulrich, R. (1999). Effects of auditory stimulus intensity on response force in simple, go/no-go, and choice RT tasks. *Perception & Psychophysics*, 61(1), 107–119. <https://doi.org/10.3758/BF03211952>
- Moradi, Z. Z., Manohar, S., Duta, M., Enock, F., & Humphreys, G. W. (2018). In-group biases and oculomotor responses: Beyond simple

- approach motivation. *Experimental Brain Research*, 236(5), 1347–1355. <https://doi.org/10.1007/s00221-018-5221-7>
- Moray, N. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology*, 11(1), 56–60. <https://doi.org/10.1080/17470215908416289>
- Nobre, A. C., & van Ede, F. (2018). Anticipated moments: Temporal structure in attention. *Nature Reviews Neuroscience*, 19(1), 34–48. <https://doi.org/10.1038/nrn.2017.141>
- Orellana-Corales, G., Matschke, C., & Wesslein, A.-K. (2021). The impact of newly self-associated pictorial and letter-based stimuli in attention holding. *Attention, Perception, & Psychophysics*. <https://doi.org/10.3758/s13414-021-02367-1>
- Otto, T. U., & Mamassian, P. (2012). Noise and correlations in parallel perceptual decision making. *Current Biology*, 22(15), 1391–1396. <https://doi.org/10.1016/j.cub.2012.05.031>
- Paquette, S., Peretz, I., & Belin, P. (2013). The “musical emotional bursts”: A validated set of musical affect bursts to investigate auditory affective processing. *Frontiers in Psychology*, 4, 509. <https://doi.org/10.3389/fpsyg.2013.00509>
- Payne, S., Tsakiris, M., & Maister, L. (2017). Can the self become another? Investigating the effects of self-association with a new facial identity. *Quarterly Journal of Experimental Psychology*, 70(6), 1085–1097. <https://doi.org/10.1080/17470218.2015.1137329>
- Payne, B., Lavan, N., Knight, S., & McGettigan, C. (2020). Perceptual prioritization of self-associated voices. *British Journal of Psychology*. <https://doi.org/10.1111/bjop.12479>
- Posner, M. I., Nissen, M. J., & Klein, R. M. (1976). Visual dominance: An information-processing account of its origins and significance. *Psychological Review*, 83(2), 157–171. <https://doi.org/10.1037/0033-295X.83.2.157>
- Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin*, 86(3), 446–461. <https://doi.org/10.1037/0033-2909.86.3.446>
- Rogers, T. B., Kuiper, N. A., & Kirker, W. S. (1977). Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology*, 35(9), 677–688. <https://doi.org/10.1037/0022-3514.35.9.677>
- Schäfer, S., Wentura, D., & Frings, C. (2015). Self-prioritization beyond perception. *Experimental Psychology*, 62(6), 415–425. <https://doi.org/10.1027/1618-3169/a000307>
- Schäfer, S., Frings, C., & Wentura, D. (2016a). About the composition of self-relevance: Conjunctions not features are bound to the self. *Psychonomic Bulletin & Review*, 23(3), 887–892. <https://doi.org/10.3758/s13423-015-0953-x>
- Schäfer, S., Wesslein, A. K., Spence, C., Wentura, D., & Frings, C. (2016b). Self-prioritization in vision, audition, and touch. *Experimental Brain Research*, 234(8), 2141–2150. <https://doi.org/10.1007/s00221-016-4616-6>
- Schäfer, S., Wentura, D., & Frings, C. (2020a). Creating a network of importance: The particular effects of self-relevance on stimulus processing. *Attention, Perception, & Psychophysics*, 82, 3750–3766. <https://doi.org/10.3758/s13414-020-02070-7>
- Schäfer, S., Wesslein, A.-K., Spence, C., & Frings, C. (2020b). When self-prioritization crosses the senses: Crossmodal self-prioritization demonstrated between vision and touch. *British Journal of Psychology*, 112, 573–584. <https://doi.org/10.1111/bjop.12483>
- Shaw, L. H., Freedman, E. G., Crosse, M. J., Nicholas, E., Chen, A. M., Braiman, M. S., Molholm, S., & Foxe, J. J. (2020). Operating in a multisensory context: assessing the interplay between multisensory reaction time facilitation and inter-sensory task-switching effects. *Neuroscience*, 436, 122–135. <https://doi.org/10.1016/j.neuroscience.2020.04.013>
- Sinnett, S., Spence, C., & Soto-Faraco, S. (2007). Visual dominance and attention: The Colavita effect revisited. *Perception & Psychophysics*, 69(5), 673–686. <https://doi.org/10.3758/BF03193770>
- Sinnett, S., Soto-Faraco, S., & Spence, C. (2008). The co-occurrence of multisensory competition and facilitation. *Acta Psychologica*, 128(1), 153–161. <https://doi.org/10.1016/j.actpsy.2007.12.002>
- Spence, C., & Driver, J. (1999). A new approach to the design of multimodal warning signals. In D. Harris (Ed.), *Engineering psychology and cognitive ergonomics, Vol. 4: Job design, product design and human-computer interaction* (pp. 455–461). Ashgate Publishing.
- Spence, C., Nicholls, M. E. R., & Driver, J. (2001). The cost of expecting events in the wrong sensory modality. *Perception & Psychophysics*, 63, 330–336. <https://doi.org/10.3758/BF03194473>
- St Germain, L., Smith, V., Maslovat, D., & Carlsen, A. (2020). Increased auditory stimulus intensity results in an earlier and faster rise in corticospinal excitability. *Brain Research*, 1727, 146559. <https://doi.org/10.1016/j.brainres.2019.146559>
- Stein, T., Siebold, A., & van Zoest, W. (2016). Testing the idea of privileged awareness of self-relevant information. *Journal of Experimental Psychology: Human Perception and Performance*, 42(3), 303–307. <https://doi.org/10.1037/xhp0000197>
- Stokes, M., Thompson, R., Nobre, A. C., & Duncan, J. (2009). Shape-specific preparatory activity mediates attention to targets in human visual cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 106(46), 19569–19574. <https://doi.org/10.1073/pnas.0905306106>
- Stolte, M., Humphreys, G., Yankouskaya, A., & Sui, J. (2017). Dissociating biases towards the self and positive emotion. *Quarterly Journal of Experimental Psychology*, 29, 1–12. <https://doi.org/10.1080/17470218.2015.1101477>
- Stolte, M., Spence, C., & Barutcu, A. (2021). Multisensory perceptual biases for social and reward associations. *Frontiers in Psychology*, 12, 640684. <https://doi.org/10.3389/fpsyg.2021.640684>
- Sui, J., & Humphreys, G. W. (2015). The interaction between self-Bias and reward: evidence for common and distinct processes. *Quarterly Journal of Experimental Psychology*, 68(10), 1952–1964. <https://doi.org/10.1080/17470218.2015.1023207>
- Sui, J., & Humphreys, G. W. (2017a). Aging enhances cognitive biases to friends but not the self. *Psychonomic Bulletin & Review*, 24(6), 2021–2030. <https://doi.org/10.3758/s13423-017-1264-1>
- Sui, J., & Humphreys, G. W. (2017b). The ubiquitous self: What the properties of self-bias tell us about the self. *Annals of the New York Academy of Sciences*, 1396(1), 222–235. <https://doi.org/10.1111/nyas.13197>
- Sui, J., He, X., & Humphreys, G. W. (2012). Perceptual effects of social salience: Evidence from self-prioritization effects on perceptual matching. *Journal of Experimental Psychology: Human Perception and Performance*, 38(5), 1105–1117. <https://doi.org/10.1037/a0029792>
- Sui, J., Rotshtein, P., & Humphreys, G. W. (2013). Coupling social attention to the self forms a network for personal significance. *Proceedings of the National Academy of Sciences of the United States of America*, 110(19), 7607–7612. <https://doi.org/10.1073/pnas.1221862110>
- Symons, C. S., & Johnson, B. T. (1997). The self-reference effect in memory: A metaanalysis. *Psychological Bulletin*, 121(3), 371–394. <https://doi.org/10.1037/0033-2909.121.3.371>
- Talsma, D. (2015). Predictive coding and multisensory integration: An attentional account of the multisensory mind. *Frontiers in Integrative Neuroscience*, 9(19), 1–13. <https://doi.org/10.3389/fnint.2015.00019>
- Talsma, D., Senkowski, D., Soto-Faraco, S., & Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends in Cognitive Sciences*, 14(9), 400–410. <https://doi.org/10.1016/j.tics.2010.06.008>

- Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Basil Blackwell.
- Verma, A., Jain, A., & Srinivasan, N. (2021). Yes! I love my mother as much as myself: Self- and mother-association effects in an Indian sample. *Quarterly Journal of Experimental Psychology*, 174702182111033118. <https://doi.org/10.1177/174702182111033118>
- Wade, G. L., & Vickery, T. J. (2018). Target self-relevance speeds visual search responses but does not improve search efficiency. *Visual Cognition*, 26(8), 563–582. <https://doi.org/10.1080/13506285.2018.1520377>
- Wood, N., & Cowan, N. (1995). The cocktail party phenomenon revisited: How frequent are attention shifts to one's name in an irrelevant auditory channel? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(1), 255–260. <https://doi.org/10.1037//0278-7393.21.1.255>
- Woźniak, M., & Knoblich, G. (2019). Self-prioritization of fully unfamiliar stimuli. *Quarterly Journal of Experimental Psychology*, 72(8), 2110–2120. <https://doi.org/10.1177/1747021819832981>
- Woźniak, M., & Knoblich, G. (2021). Self-prioritization depends on assumed task-relevance of self-association. *Psychological Research*, 1–16. <https://doi.org/10.1007/s00426-021-01584-5>
- Woźniak, M., Kourtis, D., & Knoblich, G. (2018). Prioritization of arbitrary faces associated to self: An EEG study. *PLoS One*, 13(1), 1–22. <https://doi.org/10.1371/journal.pone.0190679>
- Wundt, W. (1910). *Principles of physiological psychology* (Vol. 1, 2nd ed.). Swan Sonnenschein & Co. <https://doi.org/10.1037/12381-000>
- Yankouskaya, A., Humphreys, G., Stolte, M., Stokes, M., Moradi, Z., & Sui, J. (2017). An anterior–posterior axis within the ventromedial prefrontal cortex separates self and reward. *Social Cognitive and Affective Neuroscience*, 12(12), 1859–1868. <https://doi.org/10.1093/scan/nsx112>
- Yin, S., Sui, J., Chiu, Y.-C., Chen, A., & Egner, T. (2019). Automatic prioritization of self-referential stimuli in working memory. *Psychological Science*, 30(3), 415–423. <https://doi.org/10.1177/0956797618818483>
- Zuanazzi, A., & Noppeney, U. (2019). Distinct neural mechanisms of spatial attention and expectation guide perceptual inference in a multisensory world. *Journal of Neuroscience*, 39(12), 2301–2312. <https://doi.org/10.1523/JNEUROSCI.2873-18.2019>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.