Social relevance modulates multisensory integration

Meike Scheller[1,2], Jie Sui[1]

[1]School of Psychology, University of Aberdeen

[2]Department of Psychology, Durham University

Open Science Disclosures: The study materials are available on the Open Science Framework (OSF), and can be accessed here: https://osf.io/z2gtm/

The material includes:

- Pre-Registration, including analysis plan
- Stimuli and scripts necessary to run the experiment in Inquisit
- Data and analysis scripts

The control experiment (experiment 3) has not been pre-registered. Any further deviations from pre-registration are outlined in section S3 of the supplementary material.

Author Note: We have no conflicts of interest to disclose.

Correspondence concerning this article should be addressed to Meike Scheller (meike.scheller@durham.ac.uk), Department of Psychology, University of Durham, DH1 3LE, Durham, United Kingdom.

Abstract

When interacting with the environment, humans exhibit robust biases towards information that pertains to themselves: Self-relevant information is processed faster and yields more accurate responses than information linked to others. Recent studies have shown that simple social associations can lead to the instant deployment of this benefit in the processing of abstract stimuli. However, how self-prioritization evolves across the processing hierarchy has been a subject of intense debate. Furthermore, there is little empirical evidence about the functional efficiency of social relevance in natural environments in which information is present across multiple senses. Across three experiments (each *n* = 40), the present study shows that self-prioritization effects (1) can arise in simple audio-visual numerosity judgements, (2) can be efficiently deployed across the senses by funnelling perception towards self-relevant information in the more reliable sensory modality, and (3) modulates the integration of auditory and visual information into a multisensory representation. Taken together, the present findings suggest that social salience can influence multisensory processing at both perceptual and post-perceptual stages via early attentional modulations of sensory integration and later, task-dependent attentional control.

Significance statement

This study provides evidence that self-relevance of abstract information (temporal event numerosity) leads to changes in multisensory perception. Establishing self-relevance via social associations of oneself or another person with two visual event numerosities leads to a relative performance facilitation for self-related information that transfers to auditory and multisensory contexts. Furthermore, in relation to other-associated information, self-association with event numerosities modulates audio-visual integration by increasing sensitivity and decreasing bias in the fusion illusion. Together, the present findings suggest that social salience can influence multisensory processing via attentional modulations.

Keywords: Self-prioritization, Multisensory integration, Cross-modal processing, Social relevance, Information processing

1. Introduction

We live in a multisensory world in which concurrent information that reaches us through our different senses feeds into a coherent percept of our environment, forming the basis of our experiences and allowing us to create a stable viewpoint through which we perceive and interact with the world around us. This viewpoint, which lets us experience our environment with spatial and temporal continuity, can also be described as a fundamental form of self-representation (Damasio, 2003; Northoff, 2016) that allows us to distinguish our self from our environment and others. While some aspects of our self-representation are likely innate and highly evolutionarily conserved (Northoff & Panksepp, 2008; Panksepp & Northoff, 2009), others develop only with repeated experience throughout childhood (Cowie et al., 2018; Nava et al., 2018) and even remain plastic into adulthood (Cardini et al., 2013; Ehrsson, 2020; Tajadura-Jiménez et al., 2012; Tsakiris, 2010, 2017). For example, a large body of research suggests that our physical self-representation, operationalized through body ownership, is malleable to multisensory experiences: the simultaneous stroking of a (seen) fake hand and the (felt) own hand can lead to the experience of the rubber-hand-illusion, whereby we experience body ownership over the fake hand as a result of mere visuo-tactile perceptual coherence (Botvinick & Cohen, 1998; Ehrsson, 2020; Samad et al., 2015). Moreover, recognition of own body parts alters the degree to which body ownership can be changed through multisensory coherence (O' Dowd et al., 2020; O'Dowd & Newell, 2020; Pyasik et al., 2020; Tsakiris, 2008). These studies demonstrate that multisensory redundancy provides a basis upon which our (physical) self-image is established, and that familiarity and repeated experience with this image can determine its flexibility to incorporate novel information into its representation. On the contrary, direct effects of our internal self-representation on multisensory processing have been comparably less investigated. Most work exploring the role of the self on multisensory perception has been confined to the effects of peripersonal space (Bassolino et al., 2010; Noel et al., 2018; Pellencin et al., 2018; Serino, 2019), which links the experience of bodily self-consciousness (Noel et al., 2015, 2018) and social closeness

(Teneggi et al., 2013) to changes in multisensory processing of information presented in close proximity to the own body.

At the same time, experimental research in the area of social cognition has operationalized the self in terms of the degree of self-relatedness with external stimuli, showing how the attachment of social relevance to arbitrary, external information gives rise to instant processing advantages proportional with the social proximity to the self (Sui et al., 2012). That is, by associating information with oneself, either through ownership (Golubickis et al., 2018; Truong et al., 2017) or the integration of information into a stable, internal representation of the self (Falbén et al., 2020; Hu et al., 2016; Sui et al., 2012; Woźniak & Knoblich, 2019), performance is enhanced, reflected in more accurate, sensitive, and accelerated responses towards this information compared to information associated with other individuals (Humphreys & Sui, 2016; Sui et al., 2012; Sui & Humphreys, 2015b). This Self-Prioritization Effect (SPE) has been robustly evidenced across different tasks and types of stimuli (Hu et al., 2016; Northoff, 2016; Northoff & Hayes, 2011; Sui & Humphreys, 2017; Woźniak & Knoblich, 2019). Its evolutionary perseverance (Northoff & Panksepp, 2008; Panksepp & Northoff, 2009) specificity (Constable, Rajsic, et al., 2019; Schäfer, Frings, et al., 2016), and contextual flexibility (Caughey et al., 2020; Falbén, Golubickis, Wischerath, et al., 2020) suggest that self-prioritization plays an adaptive role in information processing.

1.1. Does self-relevance modulate ‘perception’? The need for a clearer definition.

While the functional role of self-prioritization is well-placed in the adaptive context, there is less consensus regarding its mechanistic generation. While accumulating neuropsychological and behavioural evidence suggests that self-related behavioural facilitation arises across multiple stages of information processing, the contribution of the different processing stages, particularly perceptual processing (Macrae et al., 2017b; Reuther & Chakravarthi, 2017; Siebold et al., 2015; T. Stein et al., 2016; Sui et al., 2012), is still debated (Caughey et al., 2020; Constable et al., 2011; Constable, Welsh, et al., 2019; Conway & Pleydell-Pearce, 2000;

Cunningham et al., 2008; Desebrock & Spence, 2021; Frings & Wentura, 2014; Golubickis et al., 2021; Sui & Humphreys, 2013).

There have been two crucial limitations in discerning perceptual contributions to SPE generation, in contrast to other, higher levels of processing. Firstly, most studies that aimed to establish the effects of self-association on perception relied on the perceptual matching paradigm (e.g. Moradi et al., 2015; B. Payne et al., 2020; S. Payne et al., 2017; Reuther & Chakravarthi, 2017; Sui et al., 2012). However, the paradigm does not, in itself, isolate perceptual processes. In this paradigm, social identities are associated with arbitrary perceptual stimuli, which are then judged on their categorical congruency with a label that denotes one of the social identities (but see Lee et al., 2021). Performance on those tasks does not depend on a single, but on multiple processing stages involving perception, memory, and decision-making (Sui & Humphreys, 2015b). Therefore, the question of perceptual influences on self-prioritization cannot be answered with this paradigm alone, but needs to be addressed by decomposing the perceptual domain and targeting specific perceptual sub-processes (see e.g. Macrae et al., 2017; Sui et al., 2012, Exp 4a; Truong et al., 2017).

Secondly, and even more importantly, there is lacking consensus as to what perception is defined as. This lack of definition leads to strong differences in assumptions that authors make about ‘perception’, for instance, whether it is penetrable by higher level concepts such as semantic meaning, knowledge about statistical regularities, or semantically-informed attentional modulations at all, such as in the case of self-relevance (Cermeño-Aínsa, 2020; Çukur et al., 2013; Firestone & Scholl, 2015; Newen & Vetter, 2017). While some authors’ definition of perception tends to be limited to the very early decoding stages of sensory (mostly unisensory) information, this definition has been argued to neglect important aspects that influence how individuals perceive information in real-world environments (Ghazanfar & Schroeder, 2006; Matusz et al., 2018). We adopt a broader definition that largely aligns with the views of active inference and the Bayesian brain (Friston, 2010; Hohwy, 2017; Knill & Pouget, 2004; Otten et al., 2017), in which perception is considered a multi-stage, probabilistic process. Our definition understands perception as the translation from peripheral sensory input into a coherent internal representation (percept) of the external environment. Notably, this necessitates an interplay of processes that do not only progress in a strict bottom-up fashion, but that can be modulated by prior experience, top-down attentional control, as well as task-demands (Barbot & Carrasco, 2017; Cermeño-Aínsa, 2020; Den Ouden et al., 2012; Jepma et al., 2018; Liu & Sui, 2016; Mathes et al., 2006; Newen & Vetter, 2017; Odegaard & Shams, 2016; Otten et al., 2017; Picard & Friston, 2014; Ting et al., 2015; Wahn & König, 2017; Wang et al., 2019; Wutz et al., 2018).

1.2. Functional efficiency of cross-modal transfer

As outlined above, information is usually present across multiple sensory modalities, hence, the integration of redundant feature information (e.g. size, location, temporal dynamics, numerosity) within and across the different senses plays a core role in translating peripheral sensory inputs into a coherent percept (Munoz & Blumstein, 2012; Stein, 2012). However, the vast majority of studies investigated self-prioritization in unisensory, visual tasks, with only few studies providing further evidence for the beneficial effects of self-relevance on auditory or tactile processing (B. Payne et al., 2020; Schäfer, Wesslein, et al., 2016).

Recent evidence, however, suggests that newly-learnt self-relevance operates at a modality-independent, supramodal[1] level. In their study, Schäfer and colleagues (2020) showed that the beneficial effects of arbitrarily assigned self-relevance, associated with the temporal sequence of tactile taps, can transfer to the visual modality. That is, while social associations between three different rhythms and three social identities (self, friend, stranger) were trained in the tactile modality, they led to measurable SPEs in the visual modality. These findings suggest that the benefits of self-association can automatically be deployed across

[1]supramodal, or multimodal, refers to feature representations that can be accessed from more than one sense, without changing its inherent meaning. That is, stimulus features such as spatial location, movement, intensity, timing, duration, or numerosity have the same perceptual and semantic meaning, independently of whether they are accessed via vision, audition, or touch. The use of supramodal in the current context further aims at avoiding to fuel the debate surrounding the modal vs amodal dichotomy of conceptual representations, which lies outside the scope of the present study, and seeks to reduce the conflation of conceptual amodal with supramodal attributes (Barsalou, 2016; Haimovici, 2018; Michel, 2020).

sensory modalities, however, how self-relevance regulates the perceptual mechanisms that lead to their benefit remains elusive. Furthermore, these findings pose the functional question of whether the conceptual coupling of social relevance with supramodal stimulus features is used efficiently across different contexts, or whether it is more strongly influenced by the learned, sensory memory trace via which the association was formed. In other words, one may ask whether the SPE is expressed more strongly in the sense in which the association was learned, or the sense that is more reliable for the task at hand. Given that any effects of self-relevance on perceptual processing would likely result from attentional modulations (Carrasco, 2009; Humphreys & Sui, 2016; Rohe & Noppeney, 2018; Störmer et al., 2009; Sui & Rotshtein, 2019), one would assume that functionally efficient self-prioritization is expressed differently across the senses, depending on their respective sensory reliability.

1.3. Present study: Merging paradigms to assess how multisensory integration contributes to self-prioritization

In order to address these questions, the present study moved beyond the simple perceptual matching paradigm and a single sensory modality, and assessed how efficiently the self-relevance benefit was deployed across the senses (functional aspect), and whether it affected the integration of multisensory information (mechanistic aspect). We used an audio-visual temporal numerosity judgement task, an adaptation of the sound-induced flash-illusion (Andersen et al., 2004; Shams et al., 2000, 2002; see Hirst et al., 2020 and Keil, 2020 for recent reviews) which allows to dissociate the influence of auditory and visual information onto the final established percept. In a recent study, Wahn and colleagues (2020) found that multisensory performance on the sound-induced flash-illusion is altered by the presence of another individual, suggesting malleability of the illusion to social attention.

In the present task, participants observed either two or three events in either a single sensory modality (visual flashes, auditory beeps), or in both modalities together, and were asked to report how many events they perceived. On all but the auditory-only trials, participants were required to report the number of flashes they perceived. Importantly, due to structural and functional differences in the auditory and visual sensory systems, audition has a higher temporal resolution than vision, making it the more reliable sense for estimating event numerosity in a fast sequence of events (Lukas et al., 2014; Morein-Zamir et al., 2003; Recanzone, 2003; Scheller et al., 2018; Shams et al., 2002; Shimojo et al., 2001). This difference in reliability between the senses typically leads to a strong influence of audition on vision. When auditory and visual event numerosity is put into conflict, the higher auditory reliability causes illusory visual percepts, whereby participants either report less visual events than were presented (fusion) or report a further visual event that was not presented (fission; Andersen et al., 2004; Shams et al., 2002). The susceptibility to the illusion is often used as an index for the integration of auditory and visual information (Keil, 2020; Shams et al., 2005). By associating the two event numerosities with two social identities (self, other) we introduced social salience to the external stimulus. In order to allow experimental control over the decisional dimension of interest (number vs social, see methods for more details) the event numerosity estimation task was embedded in a sequential matching design.

1.4. Research question and hypotheses

The overall question the present study addressed was whether and how self-relevance affects cross-modal processing. This question was considered at three levels:

(1) Firstly, in order to confirm whether social association with event numerosity led to a self-prioritization effect (SPE), we compared accuracy[2] between self- and other-associated event numerosities in the bimodal congruent conditions, in which redundant sensory evidence was present. Our first hypothesis (H1) stated that numerosity information would be processed more accurately when it was associated with the self, compared to when it was associated with the other social identity.

[2] All main analyses were focussed on accuracy rather than response times (RT) as accurate performance is directly linked to the numerosity percept that was subject to the social manipulation, and might therefore be more indicative of modulatory effects of endogenous attention on perception (Prinzmetal et al., 2005, 2009). RTs may reflect different, but complimentary, attentional processes, and more detail on RT-based analyses are provided in supplementary material S2.3, offering further information on the mechanistic nature of SPEs.

(2) To probe the efficient deployment of self-relevance benefits across the senses, we assessed the presence of SPEs in the visual condition, in which the social associations were learned, and in the auditory condition, which was more reliable for the task, separately. Based on findings from (Schäfer et al., 2020), our second hypothesis (H2) was that cross-modal transfer would lead to a measurable accuracy-benefit of self-association (compared to other-association) in the auditory condition. However, these findings leave open as to whether such self-association benefit would also emerge in the visual condition. (H2a) If an SPE was present and stronger in the visual than the auditory condition, this would suggest that the learned, visual template mediates the cross-modally transferred auditory SPE. (H2b) If an SPE was absent or smaller in the visual than auditory condition, this would suggest that the learned association between social relevance and stimulus feature can be deployed efficiently based on the sensory appropriateness (sensory reliability) for the task.

(3) Thirdly, in order to assess the effects of self-association on multisensory integration (perceptual mechanisms of self-prioritization), we determined changes in the susceptibility towards the fission and fusion illusion in the bimodal incongruent conditions following social association. Our third hypothesis (H3) stated that social association would alter the susceptibility to the illusions. Depending on the possibility of cross-modal transfer of self-prioritization (i.e. H2) the directionality of the effect in H3 would differ. Effects of self-association on multisensory integration were followed up with a signal detection approach to determine whether socially induced changes in illusion susceptibility were reflected in changes of perceptual sensitivity or perceptual bias.

For all sensory conditions, each participant completed a baseline (number) task prior to social association, which served as a within-participant control that ascertained that differences between self- and other-associated information resulted from social association, and did not merely reflect individual pre-existing biases and perceptual differences of event numerosities. In the baseline task, judgements were based on event numerosity instead of social identity.

Experiment 1 formed the original experiment testing the first two hypotheses, while experiment 2 formed the replication experiment for these two hypotheses. Hypothesis 3 required the splitting of task data into separate illusion conditions, thereby halving the number of trials for each analysis. Hence, data from the fission and fusion trials was pooled across experiment 1 and 2 to provide sufficient power (see section 4. 'Multisensory integration: Fission and Fusion' for more details[3]). Experiment 3 formed a control experiment, whereby social associations were replaced by associations with regular household items. This way, potential effects that result from task demands that are unrelated to social associations could be detected. Note that for this experiment sample size was not doubled to assess hypothesis 3, limiting the power of this particular analysis.

1.5. Open Science Statement

Study hypotheses, experimental methods and analyses have been pre-registered on the Open Science Framework (Scheller & Sui, 2021). The control experiment (experiment 3) has not been pre-registered, but closely follows the same principles and analyses pipelines as experiment 1 and 2. Any further deviations from pre-registration are outlined in section S3 of the supplementary material. Experimental tasks and stimuli necessary for replication, experimental data, analysis scripts, as well as data from an initial pilot experiment, conducted to estimate effect sizes and perform sample size estimation, are stored in the relevant components of the OSF project-specific repository (https://osf.io/z2gtm/).

2. Experiment 1

2.1 Methods

2.1.1. Participants

In Experiment 1, 40 participants (25 female, age: $M \pm SD$ = 28.6±5.1years) were included in the study and were randomly allocated to one of two groups (n_1 = 20; n_2 = 20). A

[3] For completeness, results are reported for each experiment individually and for experiment 1 and 2 combined.

further 6 participants were initially recruited but excluded due to low accuracy (< 65%) in the bimodal congruent or either of the unisensory conditions at baseline (number matching task). All participants were reported to have normal or corrected-to-normal vision and audition.

Sample size estimation was conducted via Monte Carlo simulations to determine power for a generalized mixed effect model on accuracy data using the interaction of task type (number vs social) and associated identity (self vs other - and the respective number matched for each participant in the number condition) as predictors, with participant ID entering the model as a random effect. Pilot data from 11 naïve participants showed an estimated effect size of 0.842 for the bimodal congruent condition (hypothesis 1). To prevent effect overestimation and gain a more conservative measure of power an interaction effect size of 0.75 was used for simulation. Based on 1000 simulations and an alpha level of 0.05, it was determined that 40 participants would allow to detect an effect of similar size with 92.1% [CI95%:90.25; 93.7%] power. We did not estimate effect sizes for illusion effects (bimodal incongruent conditions) as no consistent effect size cold be estimated for the small sample (n = 4 and n = 7 for each illusion condition, respectively). The study was approved by the University of Aberdeen Psychology Ethics Committee (#4669/2021/1).

2.1.2. Stimuli and sensory conditions

Participants were recruited online via the platform Prolific. The experiment was conducted using Millisecond's software Inqiusit Web (Millisecond, 2020) on the participant's machine to ensure reliable timing precision. Stimulus timing was adjusted for 60Hz refresh rate monitors. Participants were required to confirm monitor refresh rate and the use of headphones. Stimulus parameters were chosen and piloted to allow enhanced consistency across different presentation machines.

Unisensory events consisted of two or three successive auditory beeps or visual flashes. The visual flashes consisted of uniform light grey discs, approximately 2.47cm in diameter, presented for the duration of 17ms (1 frame) in succession for two or three times, with a 66ms (~4 frames) break between each flash, on a dark grey background. Auditory beeps consisted of sound bursts with a frequency of 3.5kHz, presented for 10ms with a break duration of 58ms. As auditory amplitude was outside the experimenters' control, pink noise was played throughout the actual task to allow for a more standardized signal-to-noise-ratio (SNR) across different user set ups. Sounds were generated in Audacity, with the auditory event amplitude being louder than the background noise (average SNR = 2.8; $SNR_{3\text{-}4kHZ}$ = 5.4). The final parameters were found suitable based on data from five additional pilot participants.

Bimodal events consisted of a combination of the visual flashes and auditory beeps that were either presented congruently or incongruently (Figure 1d). On congruent trials, visual and auditory stimuli conveyed the same event numerosity, e.g. two flashes and two beeps. On incongruent trials, visual and auditory stimuli conveyed the opposite numerosity of events, e.g. two flashes and three beeps. Experiment 1 focussed on bimodal congruent and unimodal conditions only. Bimodal incongruent trials were analysed across the two experiments as they were split into two different combinations of visual and auditory events. Audio-visual timing was aligned relative to the respective mid-latency of trials (Figure 1e). Each trial was preceded by a variable stimulus onset delay randomly drawn out of five possible delays between 17ms and 102ms.

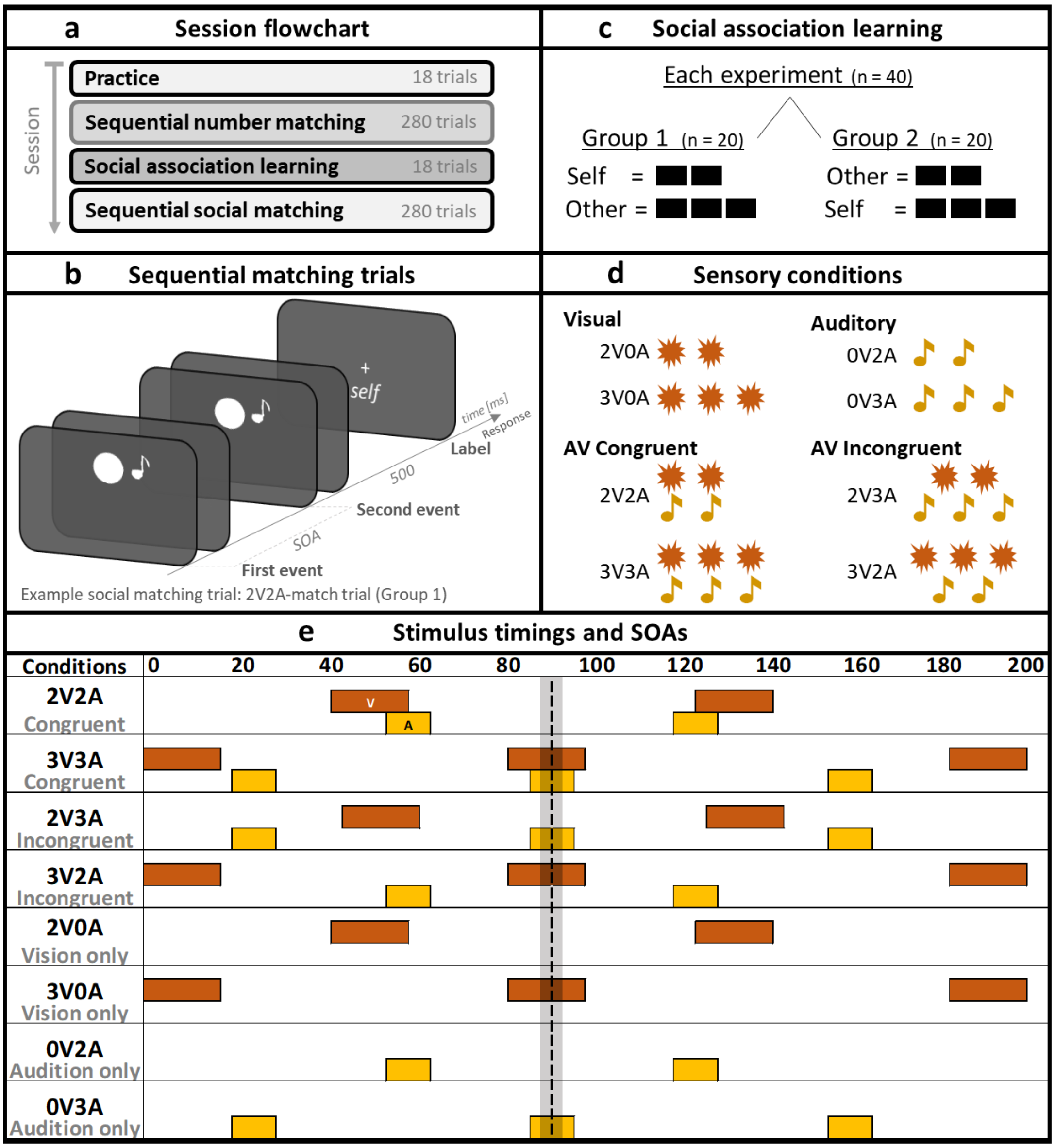

Figure 1. Experimental procedure overview: Panel a outlines the order in which participants conducted the matching tasks and association learning. Panel b illustrates an example trial of the sequential social matching task, in which a specific number of events was presented followed by a social label. Participants responded whether the label and event-association matched or mismatched. Panel c shows the associations between social identity and visual number of events that participants learned in each group. Social-numerosity associations were balanced across two groups within each experiment. Panel d outlines the different possible sensory conditions for each stimulus numerosity. V and A indicate visual and auditory events, respectively. Panel e shows stimulus timings and stimulus onset asynchronies (SOAs) for the different sensory conditions. Stimulus timing across conditions was adjusted to align with the temporal midpoint. Each stimulus was preceded by a variable onset delay to decrease stimulus predictability.

2.1.3. Procedure

The experiment consisted of two task types that both involved temporal numerosity judgements, which were based on the sound-induced flash illusion (SIFI; Shams et al., 2000, 2005, 2002; see Hirst et al., 2020 and Keil, 2020 for reviews). However, contrary to the classic SIFI paradigm which involves a single event paired with two events, the present task included two and three events in order to increase sensory uncertainty. Participants observed a number of events and had to report how many events they perceived. Trials were presented in either of the four sensory conditions (visual, auditory, bimodal congruent, bimodal incongruent) in a random order, with the frequency of trials of each condition counterbalanced across blocks.

Both task types were identical, with the difference of the dimension along which participants had to make their judgement, that is, whether the number of events matched a consecutively presented number-label ("two", "three"), or whether the number of events matched a consecutively presented, associated social identity label ("self", "other"). In order to control the decisional dimension between task types, the cross-modal numerosity judgement task was combined with a sequential matching task. That is, in the first task type (number task), which served as a baseline condition, participants had to indicate the number of events they perceived by reporting whether the number of events matched a number label that was presented with a 500ms onset delay. For the second task type, participants learned associations between the two event frequencies and themselves and another participant. For example, the participants were told that three events were associated with them, and two events were associated with one of the other participants in the experiment. In the following second task type (social task), participants had to indicate whether the number of events they perceived matched the social identity label that was presented 500ms after the offset of the event. To prevent influences of expectation, trials with matching and mismatching labels were presented equally often. However, similar to several previous studies (Sui et al., 2012; Sui & Humphreys, 2017a; Verplanken & Sui, 2019) and due to difficulties in disentangling effects related to perceptual and semantic conflict in the present paradigm the analysis focussed on matching trials only. Participants received 18 practice trials for each task type. In total, each participant took part in 16 conditions per task type: four sensory conditions (visual, auditory, bimodal congruent, bimodal incongruent) with two frequencies/social identities (two or three events, self- or other-associated), with either matching or mismatching label. Each of the bimodal conditions was repeated 25 times, and each unimodal condition was repeated 10 times per task type, leading to 280 trials per task type.

Overall, participants were assigned to one of two groups, each forming an association between a specific social identity and a specific number of events (Figure 1c). One group formed associations between the self and two events, and the other participant and three events, while the second group formed associations between the self and three events, and the other participant with two events. Sample size was equated across the groups.

2.1.4. Design

The experiment followed a full within-subject design. That is, all 40 participants completed all conditions. To test H1 and H2, the following sensory conditions were analysed separately: bimodal congruent, visual, and auditory. Within each sensory condition, social identity (or the individual number-equivalent) and task type were contrasted as within-participant factors.

2.1.5. Data analysis

To assess whether the present task allows determining self-prioritization effects (SPE) in event numerosity judgements (H1) the first analysis was focussed on bimodal congruent trials, as sensory evidence for event frequencies is highest when redundant sensory information is available. Next, (H2) to determine whether an SPE mostly presents in the sensory modality in which social association was learned (vision) or in the more reliable modality for a certain task (audition), the second analyses focussed on the visual and auditory conditions.

For each sensory condition, logistic mixed effect models were used to predict participant accuracy. In all models, task type (number and social) and social identity (self or other) were included as fixed factors and participant was entered as a random intercept effect[4]. Note that in the number task, social identity refers to the respective number of events that was associated with the given social identity for each individual (i.e. 50% of participants associated the self with two events and 50% with three events) and does not equate a specific number of events.

2.2. Results

To test hypothesis 1 and 2, the analyses were focussed on accuracy in the bimodal congruent, in the unimodal learned (visual) the unimodal reliable (audition) conditions. Across all analyses, only trials with RTs of 200ms or longer were included as responses with shorter RTs are unlikely to reflect conscious decision making. This led to the exclusion of <1% of trials in total. On average, participants' performance accuracy in the number (baseline) task in Experiment 1 was 90.5% (see Table 1). As expected based on an asymmetry in sensory reliability, auditory-only accuracy was higher than visual-only accuracy.

[4] Model specification and random effects structure selection closely followed the preregistered analysis pipeline and was based on the optimal structure from previous reports, parsimony and maximal comparability between sensory conditions. However, inclusion of within-subject factors would have allowed to select a more complex random effects structure, for which we refer the interested reader to the supplementary material S1.

Table 1: Average performance accuracy in the number (baseline) task across the four sensory conditions and all experiments. Accuracy is reported for the incongruent condition combined, as well as split for fission and fusion conditions. Lower accuracy in the incongruent conditions indicates a stronger illusion and a stronger auditory weighting.

			Number (baseline) task					
Experiment 1	n	age (mean±SD)	Congruent	Auditory	Visual	Incongruent	Fission	Fusion
All participants	40	28.59 (5.1)	92.8%	93.4%	85.4%	45.8%	55.0%	36.7%
Group 1	20	27.72 (4.4)	92.3%	91.5%	85.6%	45.4%	59.9%	30.8%
Group 2	20	28.62 (5.2)	93.3%	95.3%	85.3%	46.3%	50.0%	42.7%
Experiment 2								
All participants	40	28.90 (4.5)	90.0%	93.0%	80.9%	48.3%	55.1%	41.4%
Group 1	20	28.85 (4.4)	88.7%	93.0%	79.4%	52.5%	58.8%	46.2%
Group 2	20	29.13 (3.9)	91.0%	92.9%	81.8%	44.0%	51.4%	36.6%
Experiment 3								
All participants	40	26.08 (5.2)	86.4%	84.2%	81.4%	51.3%	58.7%	43.9%
Group 1	20	26.86 (5.0)	86.4%	83.8%	82.1%	57.5%	67.7%	47.3%
Group 2	20	25.20 (5.2)	86.4%	84.6%	80.8%	45.1%	49.7%	40.5%

In order to assess whether a self-prioritization effect was observable in the present task using social associations of event numerosity, the interaction of task type (number, social) and social identity (or identity-associated numerosity) was assessed in the bimodal congruent condition. Significant interactions between social identity-association and task type would indicate a measurable SPE, if the accuracy benefit for self-associated, compared to other-associated information is stronger in the social task than in the number-task. As expected, a significant interaction effect was present (OR(3989)=2.031, CI^{95}[1.26-3.27], p = .004, Figure 2, *top left*), showing that the self-associated number of events produced significantly higher accuracy relative to the other-associated number of events in the social condition (β = -1.284, p < .001) as compared to the baseline condition (β = -0.575, p = .007).

To test whether the self-association led to performance benefits in the modality in which social associations were learned, or in which sensory reliability was higher, interaction effects were probed separately in the visual and auditory conditions. In the visual (learned) condition, there was no significant interaction effect of task type and social identity (*OR*(1595)= 1.256, CI^{95}[0.65-2.06], p = .622, Figure 2, *top centre*). However, this condition showed enhanced accuracy performance for self-associated stimuli both in the baseline and social task types (*OR*(1595)=2.063, CI^{95}[1.36-3.13], p = .001). In the auditory (more reliable) condition, accuracy performance showed a clear self-prioritization effect indicated by an interaction of social identity and task type (*OR*(1595)= 2.83, CI^{95}[1.31 -6.12], p = .008, Figure 2, *top right*). There was no difference between self- and other-associated frequencies in the baseline condition (β = -0.097, p = .984) but a significant self-related performance benefit in the social condition (β = -1.137, p < .001).

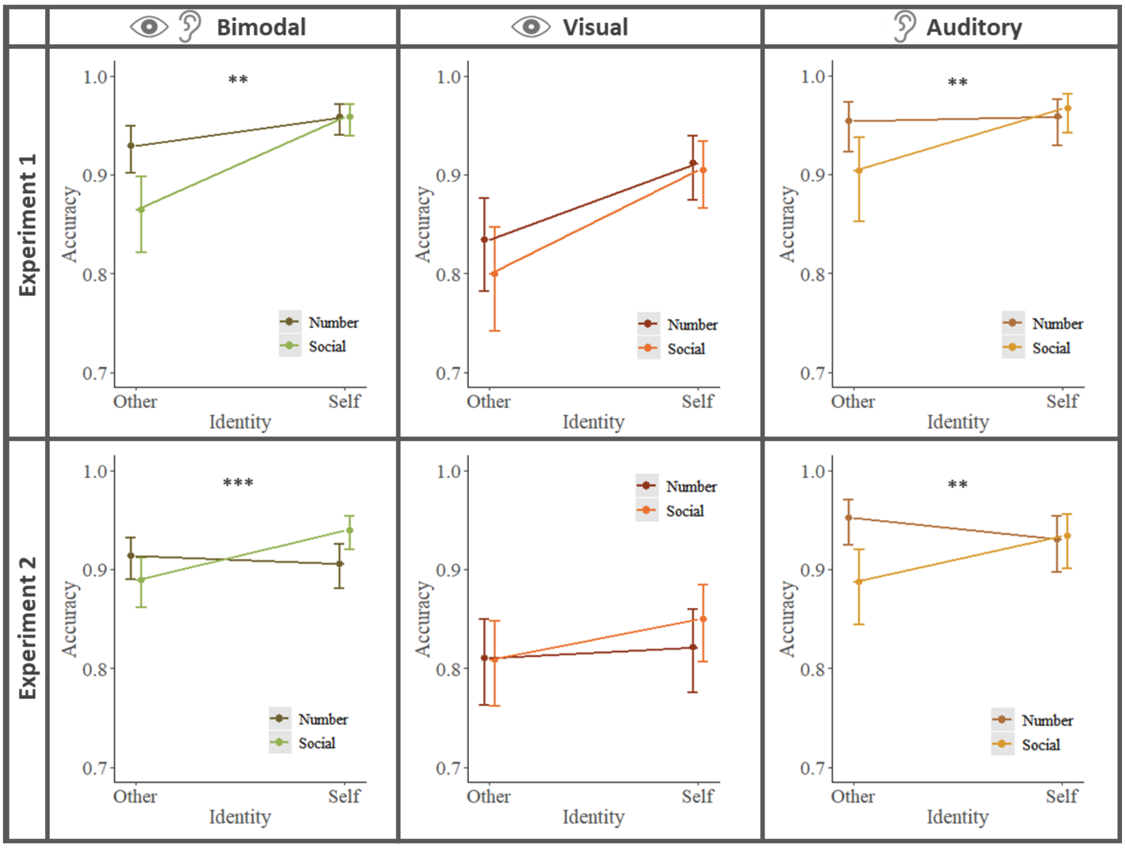

Figure 2: Marginal effects of interaction terms for accuracy estimates from the sequential matching tasks. Decision judgements were made either based on the number of events, or the associated social identities. Note that in the number task, “Other” and “Self” refer to the number of events that were associated with the respective identity in the social task. This association was evenly split between two and three events across the groups, see Figure 1c. Colors refer to the different sensory conditions (bimodal congruent: green, visual: orange, auditory: yellow) while shading indicates task judgements (number task in dark shading, social task in lighter shading). Error bars indicate standard errors. Asterisks indicate statistical significance of interaction effects. **p < .01, ***p<.001.

3. Experiment 2 (Replication)

3.1. Methods

3.1.1. Participants

For the replication experiment, another 40 naïve participants (27 female, age: *M*±*SD* = 28.9±4.5years) were recruited, and were also randomly assigned to one of two groups (n_1 = 20; n_2 = 20). Data from a further seven participants was excluded due to low accuracy (< 65%) in the bimodal congruent or either of the unisensory conditions at baseline. The same criteria that applied for Experiment 1 applied for the replication experiment.

3.1.2. Design and data analyses

The design and data analyses were the same as in experiment 1. The aim was to assess whether the findings from experiment 1 replicated in an independent sample, and to allow the analysis of illusion trials.

3.2. Results

On average, participants' performance accuracy in the number (baseline) task across bimodal congruent and unimodal conditions was lower than in Experiment 1, but remained high with 88% (see Table 1). Again, auditory-only performance exceeded visual-only performance.

To assess whether the observed benefit for the self-associated frequencies replicated in the new sample, the interaction of task type and social identity was assessed in the bimodal congruent condition. Similar to Experiment 1, the model indicated a significant interaction (OR(3991) = 2.114, CI^{95}[1.37-3.25], *p* = .001, Figure 2, *bottom left*) that was driven by the self-associated stimulus showing higher accuracy in the social condition than the other-associated stimulus (β = -0.656, p < .001). Here, there were no stimulus-based effects in the number (baseline) task (β = 0.092, p = .928).

In the visual (learned) condition, there was no significant interaction effect of task type and social identity (*OR*(1597) = 1.235, CI^{95}[0.74-2.06], *p* = .418, Figure 2, *bottom centre*). There were also no main effects of social identity (*OR*(1597) = 1.082, CI^{95}[0.76-1.54], *p* = .666) nor of task type (*OR*(1597) = 0.991, CI^{95}[0.70-1.41], *p* = .962). In the auditory (more reliable) condition, there was a significant interaction effect of task type and social identity (*OR*(1596) = 2.664, CI^{95}[1.29-5.51], *p* = .008, Figure 2, *bottom right*). Tukey-adjusted follow-up contrasts of marginal means showed that this interaction was driven by a decrease in accuracy when the number of events was associated with the other person (β = 0.929, *p* = .002), but not when it was associated with the self (β = -0.051, *p* = .997).

4. Multisensory integration: Fission and Fusion

The present analysis allowed to test hypothesis 3, that is, whether social association affects the integration of auditory and visual information. For the purpose of assessing integration, the analyses focussed on the bimodal incongruent condition, in which the number of auditory and visual events was brought into conflict. Given that both senses are not equally reliable, the condition was split into fission and fusion illusion trials.

4.1. Methods

4.1.1. Design and participants

The experiment followed a mixed between-within-subjects design. That is, all participants completed all conditions, however, the sensory condition that allowed to test H3 (bimodal incongruent) was split based on the number of flashes: 2 flashes = fission illusion, 3 flashes = fusion illusion. As social identity was associated with the number of flashes, identity-association was contrasted between participants. Similar to the experiment-specific analyses above, each participant completed both the baseline and the social task. As a reduction in trials and a change from within-to between-subject factor reduces power, we combined the samples from experiment 1 and 2 in a cross-experimental analysis, thereby yielding data from 80 participants. Hence, each social identity-association group consisted of 40 participants. For completeness, we also report the results from each individual experiment.

4.1.2. Data analyses

For each illusion condition, logistic mixed effect models were used to predict participant accuracy. Model specification was similar to experiments 1 and 2, using task type and social identity as fixed factors, with a participant-specific random intercept.

4.2. Results

4.2.1. Experiment 1

Baseline accuracy data shows that accuracy was lower in both illusion conditions compared to the congruent conditions ($t(39) = 11.7$, $p < .001$), indicating that participants experienced the illusions. Overall, participants were more susceptible to the fusion illusion than the fission illusion ($t(39) = 3.39$, $p = .002$). On fission trials, when two visual events were paired with three auditory events, no significant interaction of task type and social identity was observed ($OR(1992) = 1.129$, $CI^{95}[0.88\text{-}1.45]$, $p = .334$). In the fusion illusion condition, in which three visual events were paired with two auditory events, a significant interaction of task type and social identity was present ($OR(1990) = 1.43$, $CI^{95}[1.11\text{-}1.85]$, $p = .007$).

4.2.2. Experiment 2

Baseline accuracy data shows that accuracy was lower in both illusion conditions compared to the congruent conditions ($t(39) = 10.73$, $p < .001$), indicating that participants experienced the illusions. Overall, participants were more susceptible to the fusion illusion than the fission illusion ($t(39) = 2.89$, $p = .006$). On fission trials, when two visual events were paired with three auditory events, no significant interaction of task type and social identity was observed ($OR(1991) = 0.958$, $CI^{95}[0.75\text{-}1.23]$, $p = .732$). On fusion trials, when three visual events were paired with two auditory events, no significant interaction of task type and social identity was present ($OR(1988) = 1.13$, $CI^{95}[0.89\text{-}1.44]$, $p = .319$). This absence of an interaction effect in the present experiment may either result from less power for this analysis, or from inter-individual variability (see next section). Notably, three individuals in this experiment expressed a prominent other-bias in both bimodal congruent and incongruent conditions.

4.2.3. Cross-experimental analysis

Baseline accuracy data shows that, compared to the bimodal congruent condition, average accuracy was lower in both illusion conditions ($t(79) = 16.13$, $p < .001$), indicating that participants experienced the illusions (see Table 1). Overall, participants were more susceptible to the fusion illusion than the fission illusion ($t(79) = 4.47$, $p < .001$).

On fission trials, where two visual events were paired with three auditory events, no significant interaction of task type and social identity was observed ($OR(3983) = 1.045$, CI^{95}[0.78-1.41], $p = .771$, Figure 3, *left*). There were also no main effects of either task type ($OR(3983) = 1.085$, CI^{95}[0.88-1.34], $p = .441$) nor social identity ($OR(3983) = 1.594$, CI^{95}[0.82-3.10], $p = .17$). In the fusion illusion condition, in which three visual events were paired with two auditory events, a significant interaction of task type and social identity was present ($OR(3978) = 1.471$, CI^{95}[1.09-1.98], $p = .011$, Figure 3, *centre*). Tukey-adjusted follow-up contrasts of marginal means showed that this interaction was driven by an increase in accuracy when the number of flashes was associated with the self ($\beta = -0.626$, $p < .001$), but no significant change in accuracy following association with the other participant ($\beta = -0.241$, $p = .120$). Within individuals, social-identity-biases were correlated across the congruent and incongruent conditions ($r(80) = 0.392$, $p < .001$, Figure 3, *right*), indicating that individuals who showed stronger self-prioritization in the congruent condition also showed a stronger accuracy improvement in the respective incongruent condition. As correlation analyses were not pre-registered, see supplementary material S2.2 for further information on intra-individual correlations.

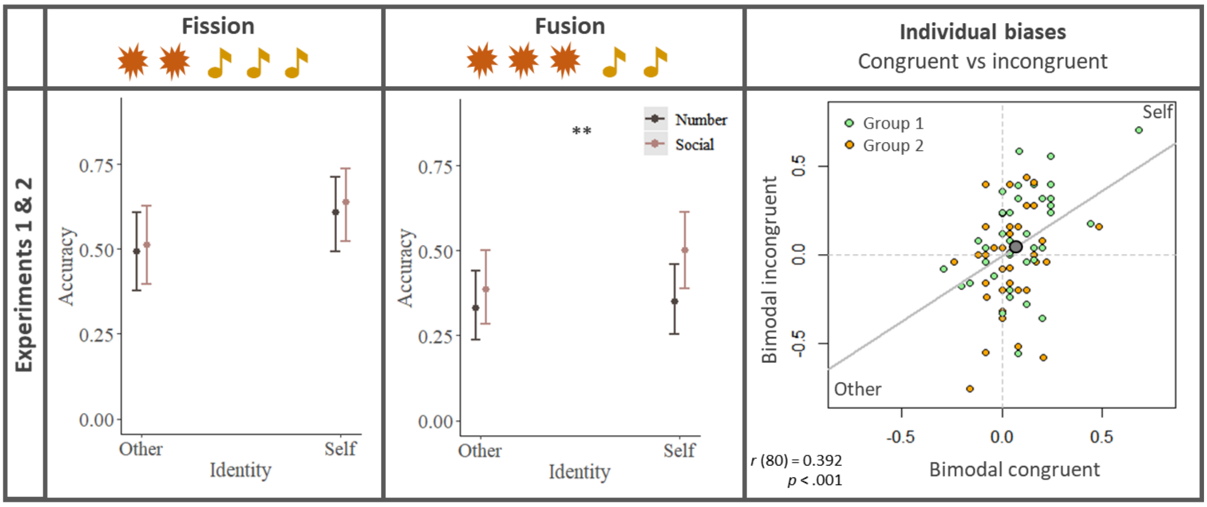

Figure 3: Marginal effects of interaction terms for accuracy estimates in the fission (two visual and three auditory events) and fusion (three visual and two auditory events) illusion conditions. Decision judgements were made either based on the number of events (Number), or the associated social identity (Social). Error bars indicate standard errors, asterisks indicate statistical significance of interaction effects. **p < .01. Right panel shows individual accuracy-bias change scores towards either the self- or other-associated numerosity in the social condition. Individual bias scores are baseline-corrected. Note that this limits improvement options when baseline accuracy is high, leading to a narrower spread in the bimodal congruent condition than in the bimodal incongruent condition. Group 1 and Group 2, both of which formed associations of the self with two events (group 1) or three events (group 2) are color-coded in green and orange, respectively. Large grey circle indicates the overall group mean, showing a deviation from the point of origin (see dashed lines). Grey line indicates line of best fit. See supplementary material S2.2 for further information.

Overall, RT analyses supported the main findings from the accuracy analyses. Across the first two experiments, RTs revealed the presence of an SPE in bimodal numerosity estimation, indicated by a speeding of responses towards self-associated compared to other-associated information in the social task ($p_{E1} < .001$; $p_{E2} < .001$). We further observed SPEs in both visual ($p < .001$) and auditory ($p < .001$) modalities in Experiment 1, and in the visual ($p < .001$), but not the auditory modality ($p = .865$) in Experiment 2. On both, fission and fusion trials, significant SPEs were measurable in RTs, with self-associated visual flash numerosity being responded to faster than other-associated flash numerosity (Fission: $p < .001$; Fusion: $p = .003$). As the effects in the unimodal conditions differ partially from those reported in the accuracy analyses, and an additional effect of self-association was measurable in the fission condition, they support the idea that RT and accuracy measures reflect different processes that complement each other in the generation of self-referenced facilitation.

Finally, follow-up analyses probed whether the effects were reflected in changes of perceptual sensitivity or perceptual bias. Importantly, as the SIFI has been extensively investigated and is evidenced to reflect a perceptual process (see Shams & Kim, 2010 for a review; also see Witt et al., 2016), its modulation via social salience is, in itself, evidence for the penetrability of perception by social salience. However, to further assess whether social salience affected multisensory integration via an increase in sensitivity of self-associated information or a decrease of auditory-induced perceptual bias, we established signal detection parameters for the detection of two and three flashes under uni- and bimodal conditions, and probed the impact of social association on these parameters. For more details on this analysis see supplementary material S4. In brief, we found that the addition of auditory stimuli enhanced the perceptual sensitivity towards three visual flashes when these were associated with the self, but not with the other. Furthermore, the bias of self-associated auditory information on other-associated visual information was increased when the other was associated with two flashes. Overall, the pre-existing asymmetry of the illusion (fusion > fission) was not significantly impacted by social association, however, the strongest measurable impact of social association on the sound-induced illusion was measured in the fusion illusion, showing an increase in perceptual bias towards self-associated three flashes.

5. Experiment 3 (Control)

The third experiment served as a control experiment, to determine whether the observed effects in experiment 1 and 2 might have resulted from other factors such as differences in task demand between number and social task, or whether the effects could largely be attributed to the social relations learned in experiment 1 and 2. To this end, the same task was used as in experiment 1 and 2, with the difference of the learned association. Here, participants were instructed to associate the two frequencies with two familiar household objects, that is, a fork and a spoon. The main objective was to assess whether prioritization effects may be present in the congruent and unimodal conditions. While the changes in integration on illusion trials were analysed as well, the sample in this analysis was half the size as that of experiment 1 and 2, for which the sample was pooled. A group-level bias towards either object (similar to an SPE in experiment 1 and 2) in either of the bimodal congruent or unisensory conditions would justify doubling the sample size for this condition, however, neither was present. While this analysis provides a direct comparison to the effects of social association on the integration of auditory and visual information, the presence or absence of effects cannot be concluded with the same power as in the previous experiments. As this is the first investigation of this type, however, it allows to explore whether any trends of group-based or individual-based biases can be observed with simple, socially-irrelevant object associations.

5.1. Methods

5.1.1. Participants

Another 40 participants (26 female, age: *M*±*SD* = 26.08±5.2years) were recruited for the control experiment, and were randomly assigned to the two association groups (n_1 = 20; n_2 = 20). A further eight participants were excluded due to low accuracy (< 65%) in the bimodal congruent or either of the unisensory conditions at baseline. The same criteria that applied to experiment 1 and 2 applied to the control experiment.

5.1.2. Design and data analyses

The design and data analyses were the same as in experiment 1 and 2, with the difference of the type of association (social identity, object type). For the bimodal incongruent condition, the design and analyses were the same as in the cross-experimental analysis, with the difference of the smaller sample size.

5.2. Results

On average, accuracy in the number (baseline) task across bimodal congruent and unimodal conditions was lower than in Experiment 1 and 2, but remained high with 84% (see Table 1). Again, accuracy was lower on the illusion trials than the bimodal congruent trials ($t(39) = 11.985$, $p < .001$), indicating that participants experienced the illusions. Also, the fusion illusion was again stronger than the fission illusion ($t(39) = 2.57$, $p = .014$).

To assess whether numerosity-association with household objects resulted in a benefit similar to social association, the interaction of task type and object type was initially assessed in the bimodal congruent condition. In contrast to experiment 1 and 2, there was no significant interaction (OR(3947) = 0.823, CI^{95}[0.57-1.19], $p = .300$, Figure 4). There were further no main effects of task type (OR(3947) = 1.079, CI^{95}[0.83-1.40], $p = .565$) nor object type (OR(3947) = 1.133, CI^{95}[0.87-1.47], $p = .347$). There were no significant interactions in neither the auditory (OR(1565) = 0.747, CI^{95}[0.42-1.34], $p = .328$) nor the visual condition (OR(1582) = 0.837, CI^{95}[0.51-1.38], $p = .485$), and no main effects of task type nor object type in either unimodal condition ($p > .360$).

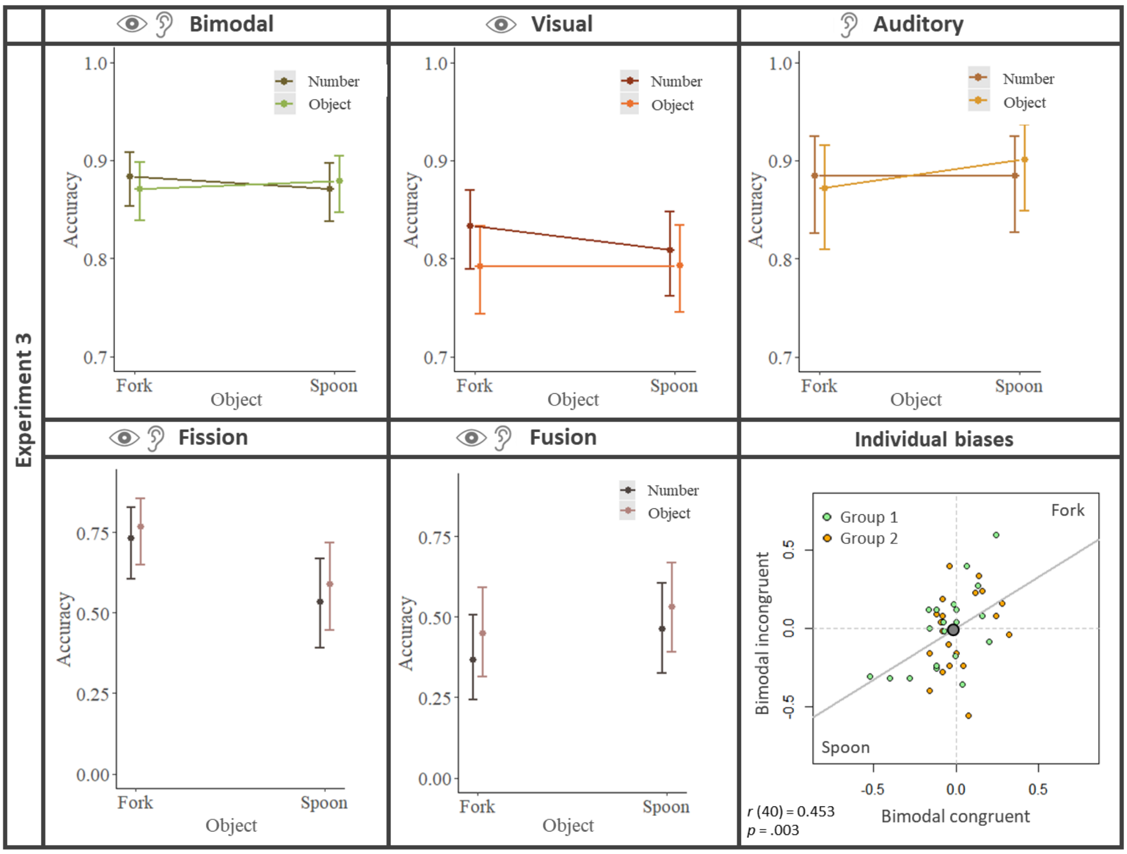

Figure 4: Marginal effects of interaction terms for accuracy estimates from the sequential matching tasks in the bimodal congruent, visual, auditory, and incongruent conditions. Decision judgements were made either based on the number of events, or the associated object. Note that in the number task, "Fork" and "Spoon" refer to the number of events that were associated with the respective item in the object task. This association was evenly split between two and three events across the groups. Error bars indicate standard errors. Colors refer to the different sensory conditions (bimodal congruent: green, visual: orange, auditory: yellow, bimodal incongruent: brown) while shading indicates task judgements (number task in dark shading, social task in lighter shading). Bottom right panel indicates individual accuracy-bias change scores towards either the fork- or spoon-associated number of flashes. Individual bias scores are baseline-corrected. See Figure 3 legend for further info and supplementary material S2.2 for discussion. Large grey circle indicates the overall group mean, showing a deviation from the point of origin (see dashed lines). Grey line indicates line of best fit.

Baseline accuracy data shows that accuracy was lower in both illusion conditions compared to the congruent conditions ($t(39) = 12.12$, $p < .001$). Overall, participants were more susceptible to the fusion illusion than the fission illusion ($t(39) = 2.47$, $p = .018$). However, in the incongruent conditions, performance was not influenced by an interaction of task type and object type in neither the fission ($OR(1967) = 0.95$, $CI^{95}[0.63\text{-}1.44]$, $p = .808$) nor the fusion illusion ($OR(1966) = 0.941$, $CI^{95}[0.63\text{-}1.41]$, $p = .765$). Individual, baseline-corrected biases towards one or the other object were correlated across the congruent and incongruent conditions ($r(40) = .453$, $p = .003$). For more information on intra-individual correlations see supplementary material S2.2. Signal detection analysis revealed no effects of object association on perceptual sensitivity or bias.

6. Discussion

In the present study, we investigated how self-relevance affects perceptual, cross-modal processing.

6.1. Social association with event numerosity leads to self-prioritization

We first assessed whether self-associated information is prioritized in numerosity estimation, using a novel combination of the sound-induced flash-illusion (Shams et al., 2000,

2002) and the sequential perceptual matching (Sui & Humphreys, 2015a; Woźniak & Hohwy, 2020) paradigms. Across two experiments, we observed that self-associated information is prioritized during perceptual numerosity estimation, as indicated by enhanced accuracy and speeding of response times (see supplementary material S.2.3 for RT analysis) towards the self- compared to the other-associated information. Such prioritized processing was not present in a control experiment in which social identities were replaced by regular household objects, suggesting that the observed effect indeed resulted from social relevance rather than task complexity or enhanced attentional and memory demand in the association condition. While these findings show that self-relevance leads to an enhanced processing of numerosity information, similar to other supramodal, temporal features (Schäfer et al., 2020; Schäfer, Wesslein, et al., 2016), they do not yet explain whether this enhancement is used efficiently and how it is modulated at perceptual processing levels. To answer these questions we assessed the deployment of the performance-benefitting association in each sensory modality alone (H2), as well as the impact of social associations on multisensory integration when both modalities received conflicting information (H3).

6.2 The deployment of self-relevance across the senses is functionally efficient

Next, we tested whether self-prioritization was predominantly expressed in the sensory modality in which it was learned and memorized (vision), or in the sensory modality that was more reliable for the task at hand (audition). Notably, the latter scenario assumes that the perception of event numerosity, which is represented as a supramodal feature (Bueti & Walsh, 2009; Nuerk et al., 2005; Piazza et al., 2006; Walsh, 2003), can be modulated by self-association via cross-modal transfer, that is, in a different modality than the one in which the association was learned. Schäfer and colleagues (2020) previously showed that such cross-modal transfer of self-related processing enhancement was possible for a different supramodal feature, i.e. temporal pattern/rhythm, between touch and vision. Here, we observed that self-prioritization was consistently expressed in the more reliable, auditory modality, but not in the visual modality. This finding suggests that self-association leads to a facilitation of perceptual processing that cannot only function beyond, but rather independently of the sensory-specific memory trace. Instead, it supports the notion that social relevance and the supramodal feature (here: numerosity) space are integrated (Sui & Humphreys, 2015b). Attentional resources can then be directed towards the specific feature property that obtained importance through self-relevance (Liu & Sui, 2016), and be deployed by the different senses across tasks in a highly efficient fashion (Driver & Spence, 2012; Humphreys & Sui, 2016; Spence & Driver, 1996), making the self-prioritization effect contextually flexible and adaptive (Scheller & Sui, 2022; Sui & Han, 2007).

It should be noted that, due to study length limitations, resulting from the high attentional demand in an online setting, both unisensory conditions included less trials than the bimodal conditions, i.e. 10 vs 25 trials per condition, respectively. Furthermore, as audition is typically more reliable for the task at hand, there was less variability in the auditory than visual unimodal condition overall. The increased variability and smaller effect size in the visual condition are therefore likely factors in explaining the failure to detect significant modulations due to self-relevance in the latter. Indeed, inspection of model parameters and effect directionality suggest a trend in line with predictions based on self-prioritization (see supplementary material S2.1. and Figure 2). However, the reduced power to detect a significant modulation of visual numerosity discrimination by social relevance does not affect the comparison of effect magnitudes between the two unisensory modalities. These indicated a stronger effect size of social salience on auditory than visual numerosity perception (OR_{aud} = 2.8, OR_{vis} = 1.2).

6.3 Self-relevance modulates multisensory integration

We further assessed whether self-relevance affects the integration of multisensory information via modulating the propensity to experience the sound-induced flash illusion (Andersen et al., 2004; Shams et al., 2000, 2002). That is, by putting auditory and visual event information into conflict we assessed the extent to which participants' perception was biased towards each numerosity, before and after event-information was associated with social identities. Here, we observed that self-relevance largely modulates the fusion, but not the fission illusion. That is, when three visual flashes were paired with two auditory beeps, participants correctly reported the three flashes more often when associated with the self, but not when associated with the other participant. This modulation did not emerge in the fission illusion, in which two flashes were paired with three beeps. In line with a previous study that conflicted the pairing of two and three audio-visual events (Andersen et al., 2004) and that showed a higher visual uncertainty for three flashes than for two flashes, we found a stronger influence of an incongruent number of beeps on three flashes (fusion) than on two flashes (fission). This suggests that, when visual discrimination was more difficult, i.e. visual uncertainty was higher, simultaneous presentation of task-irrelevant auditory information weighted more strongly into the final percept, just as would be expected based on reliability-weighting of sensory information (Ernst & Banks, 2002; Hirst et al., 2019; Shams et al., 2005). Under these conditions, self-relevance enhanced visual numerosity judgement accuracy, reducing the impact of the task-irrelevant auditory information. This was not the case when the three flashes were associated with the other social identity, nor when they were associated with either of the non-social objects. Results from our signal detection analysis support this (see supplement S4), showing that self-association led to an increase in perceptual sensitivity to three events when auditory and visual events were presented together, as well as a stronger impact of three (self-associated) sounds on the perceptual bias to perceive two (other-associated) flashes.

Taken together, our results suggest that multisensory integration is modulated by social (self-)relevance, in particular under conditions in which sensory uncertainty in the target modality is high. However, while SPEs are often stable at the group level, there is typically individual variation in the extent to which self- or other-related information is prioritized (Scheller & Sui, 2022; Stolte et al., 2017; Sui & Humphreys, 2017b). Therefore, we followed up effects in the fission and fusion illusion with an exploratory analysis (see supplementary tables S.2.1.3), in which only individuals were included who showed a self-related accuracy enhancement, relative to baseline, in the congruent condition (N = 52, n_{group1} = 23; n_{group2} = 29). In this group of participants, significant modulations of both the fission and fusion illusions were observed, with self-associated visual information being processed more accurately despite conflicting auditory information. Contrasting both illusions indicated that the magnitude of modulation in the fusion illusion was stronger than in the fission illusion, which is in line with the illusion asymmetry present in this sample (see also supplement S4), and may explain why the effect did not emerge in the overall sample. Furthermore, it is possible that those individuals that exhibited a performance enhancement for other-related information, relative to baseline (N = 17), showed a modulation on fission trials opposite to those presented by the group showing clearer self-prioritizing. A significant intra-individual correlation of socially-induced performance-enhancements between congruent and incongruent conditions supported this assumption. Interestingly, such correlation was also present in the control experiment.

The intra-individual stability across all three experiments suggests that integration of auditory and visual information may be subject to modulations through conceptual valuation or preference as well as social relevance. Indeed, previous research has shown that attention can modulate the signal reliability in the attended modality, leading to a shifting of weights in the integration process (Rohe & Noppeney, 2018; Talsma et al., 2007, 2010; Talsma & Woldorff, 2005). In instances where social associations are irrelevant, personal significance becomes an internal criterion (Nakao et al., 2010; Sui & Humphreys, 2017a). Hence, while object relevance may be individually specific, self-relevance is a feature that ubiquitously possesses stronger importance, and self-attracting biases are more prevalent than biases away from the self, both in newly-learned self-associations (Sui & Humphreys, 2017a) as well as long-term established self-associations such the own name or self-owned objects (Scheller & Sui, 2022; Tacikowski & Ehrsson, 2016) or even the space around the physical own body (Noel et al., 2015; Serino, 2019). This explains why, at the individual level, biases towards one category reflect similar performance enhancements in both bimodal tasks, while there is an average shift at group-level for the social association task only.

One may consider it likely that the increase in accuracy towards the self-associated number of flashes may result from the enhanced processing of the auditory other-associated beeps, as self-prioritization was expressed more strongly in the auditory condition. However, results from our signal detection analysis suggest that the effect in the fusion illusion was driven by an increase in sensitivity towards self-associated events ($\Delta d'_3$) and a stronger bias of the self-associated auditory information on other-associated visual numerosity (Δc_2). It is therefore worth considering the results from hypothesis 2 and 3 in light of cross-modal attentional allocation. That is, self-prioritization was expressed more strongly in the sensory modality that was more reliable for the task at hand (audition > vision), but mostly when this modality was task-relevant (Rohe & Noppeney, 2018; Shams et al., 2005; Wahn & König, 2017). However, when both sensory modalities received information, self-relevance enhanced sensitivity in the sensory modality in which the estimate had to be reported, while the more reliable modality was more efficiently suppressed (other-associated sound in $\Delta d'_3$). At the same time, self-association in the more reliable modality enhanced the perceptual bias of auditory on visual temporal numerosity estimation (Δc_3). This supports the notion that attention can flexibly modulate the integration of sensory cues (Rohe & Noppeney, 2018), depending on the sensory reliability of the modality for the stimulus quality, task-requirements, as well as direct self-relevance.

Notably, the SPE can arise via different processing streams, that is, via a salience-driven enhancement of self-associated information and attentional control for all the information at hand (Humphreys & Sui, 2016; Sui et al., 2013). While the present study cannot unequivocally distinguish between these two possibilities, the inclusion of an individual baseline task allows for informed speculation. That is, compared to the respective category in the baseline condition, we observed either an accuracy enhancement following self-association, as seen in the bimodal incongruent conditions, or an accuracy decrease following other-association, such as observed in the auditory conditions. In the bimodal congruent condition, both shifts from baseline can be observed. While it is possible that this modulation is dependent on the average baseline accuracy, that is, the modality-dependent sensory uncertainty, it might similarly suggest that the learned, visual self-association and the supramodal self-association (expressed in auditory performance) might be regulated via the different networks. In other words, while the weighted integration of audio-visual information may be influenced by social salience through an enhancement of self-related information via the ventral attention network, social modulations of auditory numerosity judgement might be achieved via the dorsal attentional control network (Humphreys & Sui, 2016; Vossel et al., 2014). The benefit observed in bimodal congruent information, on the other hand, might be the result of both networks dynamically shaping the processing advantage of self-relevance.

Taken together, these findings are well placed within the Self-Attention-Network (Humphreys & Sui, 2016) and the dual pathways of attentional modulations in multisensory integration (Tang et al., 2016). By associating arbitrary stimuli with social identities, self-relevance primes the attentional funnelling towards the stimulus feature that gained direct importance, enhancing its perceptual processing and shifting more weight to the self-associated feature during integration. The resulting supramodal representation subsequently employs executive, attentional control to influence stimulus feature estimation across modalities (Ester et al., 2016; Spagna et al., 2015), altering further processing by the associated social relations as well as task-demands. Together, these processes can lead to the functionally efficient processing effects that can be deployed across different contexts, such as task-specific decisional criteria (Caughey et al., 2020; Falbén, Golubickis, Wischerath, et al., 2020; Scheller & Sui, 2022), different stimulus materials (Scheller & Sui, 2022), and across sensory modalities (Schäfer et al., 2020; *present study*).

Given that group-based biases were present only in experiment 1 and 2, but not experiment 3, the present results cannot be explained by anything other than the social associations that participants established with the perceptual stimulus feature (numerosity). Therefore, the present study provides strong evidence for the malleability of multisensory integration by newly learned self-associations, as well as the conception that social relevance can influence information processing at multiple stages, including perception, via distinct attentional processes. It further highlights the functional significance of self-prioritization, allowing the individual to adjust different stages of information processing to flexibility adapt the prioritization of self-related information and employ their benefit across sensory modalities.

Acknowledgments

This work was supported by a grant from the Leverhulme Trust (RPG-2019-010) to JS.

References

Andersen, T. S., Tiippana, K., & Sams, M. (2004). Factors influencing audiovisual fission and fusion illusions. *Cognitive Brain Research*, *21*(3), 301–308. https://doi.org/10.1016/j.cogbrainres.2004.06.004

Barbot, A., & Carrasco, M. (2017). Attention modifies spatial resolution according to task demands. *Psychological Science*, *28*(3), 285–296. https://doi.org/10.1177/0956797616679634

Barsalou, L. W. (2016). On staying grounded and avoiding quixotic dead ends. *Psychonomic Bulletin and Review*, *23*(4), 1122–1142. https://doi.org/10.3758/s13423-016-1028-3

Bassolino, M., Serino, A., Ubaldi, S., & Làdavas, E. (2010). Everyday use of the computer mouse extends peripersonal space representation. *Neuropsychologia*, *48*(3), 803–811. https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2009.11.009

Botvinick, M., & Cohen, J. (1998). Rubber hands 'feel' touch that eyes see. *Nature 1998 391:6669*, *391*(6669), 756–756. https://doi.org/10.1038/35784

Bueti, D., & Walsh, V. (2009). The parietal cortex and the representation of time, space, number and other magnitudes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1525), 1831–1840. https://doi.org/10.1098/rstb.2009.0028

Cardini, F., Tajadura-Jiménez, A., Serino, A., & Tsakiris, M. (2013). It feels like it's me: Interpersonal multisensory stimulation enhances visual remapping of touch from other to self. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(3), 630–637. https://doi.org/10.1037/A0031049

Carrasco, M. (2009). Cross-modal attention enhances perceived contrast. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(52), 22039–22040. https://doi.org/10.1073/pnas.0913322107

Caughey, S., Falbén, J. K., Tsamadi, D., Persson, L. M., Golubickis, M., & Neil Macrae, C.

(2020). Self-prioritization during stimulus processing is not obligatory. *Psychological Research*, *85*(2), 503–508. https://doi.org/10.1007/s00426-019-01283-2

Cermeño-Aínsa, S. (2020). The cognitive penetrability of perception: A blocked debate and a tentative solution. *Consciousness and Cognition*, *77*, 102838. https://doi.org/10.1016/j.concog.2019.102838

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., Rajsic, J., Welsh, T. N., & Pratt, J. (2019). It is not in the details: Self-related shapes are rapidly classified but their features are not better remembered. *Memory and Cognition*, *47*(6), 1145–1157. https://doi.org/10.3758/s13421-019-00924-6

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*, *72*(3), 589–598. https://doi.org/10.1177/1747021818762010

Conway, M. A., & Pleydell-Pearce, C. W. (2000). The construction of autobiographical memories in the self-memory system. *Psychological Review*, *107*(2), 261–288. https://doi.org/10.1037/0033-295X.107.2.261

Cowie, D., McKenna, A., Bremner, A. J., & Aspell, J. E. (2018). The development of bodily self-consciousness: changing responses to the Full Body Illusion in childhood. *Developmental Science*, *21*(3), e12557. https://doi.org/10.1111/DESC.12557

Çukur, T., Nishimoto, S., Huth, A. G., & Gallant, J. L. (2013). Attention during natural vision warps semantic representation across the human brain. *Nature Neuroscience*, *16*(6), 763–770. https://doi.org/10.1038/nn.3381

Cunningham, S. J., Turk, D. J., Macdonald, L. M., & Neil Macrae, C. (2008). Yours or mine? Ownership and memory. *Consciousness and Cognition*, *17*(1), 312–318. https://doi.org/10.1016/j.concog.2007.04.003

Damasio, A. (2003). Mental self: The person within. *Nature 2003 423:6937*, *423*(6937), 227–227. https://doi.org/10.1038/423227a

Den Ouden, H. E. M., Kok, P., & de Lange, F. P. (2012). How prediction errors shape perception, attention, and motivation. *Frontiers in Psychology*, *3*, 548. https://doi.org/10.3389/fpsyg.2012.00548

Desebrock, C., & Spence, C. (2021). The Self-Prioritization Effect: Self-referential processing in movement highlights modulation at multiple stages. *Attention, Perception, & Psychophysics*, 1–19. https://doi.org/10.3758/s13414-021-02295-0

Driver, J., & Spence, C. (2012). Crossmodal spatial attention: Evidence from human performance. In C. Spence & J. Driver (Eds.), *Crossmodal Space and Crossmodal Attention*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198524861.003.0008

Ehrsson, H. H. (2020). Multisensory processes in body ownership. *Multisensory Perception: From Laboratory to Clinic*, 179–200. https://doi.org/10.1016/B978-0-12-812492-5.00008-5

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*(6870), 429–433. https://doi.org/10.1038/415429a

Ester, E. F., Sutterer, D. W., Serences, J. T., & Awh, E. (2016). Feature-selective attentional modulations in human frontoparietal cortex. *Journal of Neuroscience*, *36*(31), 8188–8199. https://doi.org/10.1523/JNEUROSCI.3935-15.2016

Falbén, J. K., Golubickis, M., Tamulaitis, S., Caughey, S., Tsamadi, D., Persson, L. M., Svensson, S. L., Sahraie, A., & Macrae, C. N. (2020). Self-relevance enhances evidence gathering during decision-making. *Acta Psychologica*, *209*, 103122. https://doi.org/10.1016/j.actpsy.2020.103122

Falbén, J. K., Golubickis, M., Wischerath, D., Tsamadi, D., Persson, L. M., Caughey, S., Svensson, S. L., & Macrae, C. N. (2020). It's not always about me: The effects of prior beliefs and stimulus prevalence on self–other prioritisation. *Quarterly Journal of Experimental Psychology*, *73*(9), 1466–1480. https://doi.org/10.1177/1747021820913016

Firestone, C., & Scholl, B. J. (2015). Cognition does not affect perception: Evaluating the evidence for top-down effects. *Behavioral and Brain Sciences*, *39*. https://doi.org/10.1017/S0140525X15000965

Frings, C., & Wentura, D. (2014). Self-priorization processes in action and perception. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(5), 1737–1740. https://doi.org/10.1037/a0037376

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience 2010 11:2*, *11*(2), 127–138. https://doi.org/10.1038/nrn2787

Ghazanfar, A. A., & Schroeder, C. E. (2006). Is neocortex essentially multisensory? *Trends in Cognitive Sciences*, *10*(6), 278–285. https://doi.org/10.1016/j.tics.2006.04.008

Golubickis, M., Falben, J. K., Cunningham, W. A., & Neil Macrae, C. (2018). Exploring the self-ownership effect: Separating stimulus and response biases. *Journal of Experimental Psychology: Learning Memory and Cognition*, *44*(2), 295–306. https://doi.org/10.1037/xlm0000455

Golubickis, M., Persson, L. M., Falbén, J. K., & Macrae, C. N. (2021). On stopping yourself: Self-relevance facilitates response inhibition. *Attention, Perception, and Psychophysics*, *83*(4), 1416–1423. https://doi.org/10.3758/s13414-021-02248-7

Haimovici, S. (2018). The modal—amodal distinction in the debate on conceptual format. *Philosophies*, *3*(2), 7. https://doi.org/10.3390/philosophies3020007

Helbig, H. B., & Ernst, M. O. (2008). Visual-haptic cue weighting is independent of modality-specific attention. *Journal of Vision*, *8*(1), 21. https://doi.org/10.1167/8.1.21

Hirst, R. J., McGovern, D. P., Setti, A., Shams, L., & Newell, F. N. (2020). What you see is what you hear: Twenty years of research using the Sound-Induced Flash Illusion. *Neuroscience and Biobehavioral Reviews*, *118*, 759–774. https://doi.org/10.1016/j.neubiorev.2020.09.006

Hirst, R. J., Setti, A., Kenny, R. A., & Newell, F. N. (2019). Age-related sensory decline mediates the Sound-Induced Flash Illusion: Evidence for reliability weighting models of multisensory perception. *Scientific Reports*, *9*(1), 1–12. https://doi.org/10.1038/s41598-019-55901-5

Hohwy, J. (2017). Priors in perception: Top-down modulation, Bayesian perceptual learning rate, and prediction error minimization. *Consciousness and Cognition*, *47*, 75–85. https://doi.org/10.1016/j.concog.2016.09.004

Hu, C., Di, X., Eickhoff, S. B., Zhang, M., Peng, K., Guo, H., & Sui, J. (2016). Distinct and common aspects of physical and psychological self-representation in the brain: A meta-analysis of self-bias in facial and self-referential judgements. *Neuroscience and Biobehavioral Reviews*, *61*, 197–207. https://doi.org/10.1016/j.neubiorev.2015.12.003

Humphreys, G. W., & Sui, J. (2016). Attentional control and the self: The Self-Attention Network (SAN). *Cognitive Neuroscience*, *7*(1–4), 5–17. https://doi.org/10.1080/17588928.2015.1044427

Jepma, M., Koban, L., van Doorn, J., Jones, M., & Wager, T. D. (2018). Behavioural and neural evidence for self-reinforcing expectancy effects on pain. *Nature Human Behaviour*, *2*(11), 838–855. https://doi.org/10.1038/s41562-018-0455-8

Keil, J. (2020). Double flash illusions: Current findings and future directions. *Frontiers in Neuroscience*, *14*, 298. https://doi.org/10.3389/fnins.2020.00298

Knill, D. C., & Pouget, A. (2004). The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends in Neurosciences*, *27*(12), 712–719. https://doi.org/10.1016/j.tins.2004.10.007

Lee, N. A., Martin, D., & Sui, J. (2021). A pre-existing self-referential anchor is not necessary for self-prioritisation. *Acta Psychologica*, *219*, 103362. https://doi.org/10.1016/j.actpsy.2021.103362

Liu, M., & Sui, J. (2016). The interaction between social saliency and perceptual saliency. *Quarterly Journal of Experimental Psychology*, *69*(12), 2419–2430. https://doi.org/10.1080/17470218.2015.1120330

Lukas, S., Philipp, A. M., & Koch, I. (2014). Crossmodal attention switching: Auditory dominance in temporal discrimination tasks. *Acta Psychologica*, *153*, 139–146. https://doi.org/10.1016/j.actpsy.2014.10.003

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

Mathes, B., Trenner, D., & Fahle, M. (2006). The electrophysiological correlate of contour integration is modulated by task demands. *Brain Research*, *1114*(1), 98–112. https://doi.org/10.1016/j.brainres.2006.07.068

Matusz, P. J., Dikker, S., Huth, A. G., & Perrodin, C. (2018). Are we ready for real-world neuroscience? *Journal of Cognitive Neuroscience*, *31*(3), 327–338. https://doi.org/10.1162/jocn_e_01276

Michel, C. (2020). Overcoming the modal/amodal dichotomy of concepts. *Phenomenology and the Cognitive Sciences*, 1–23. https://doi.org/10.1007/s11097-020-09678-y

Millisecond. (2020). *Inquisit* (6.4.0). Millisecond. https://www.millisecond.com

Moradi, Z., Sui, J., Hewstone, M., & Humphreys, G. W. (2015). In-group modulation of perceptual matching. *Psychonomic Bulletin and Review*, *22*(5), 1255–1277. https://doi.org/10.3758/s13423-014-0798-8

Morein-Zamir, S., Soto-Faraco, S., & Kingstone, A. (2003). Auditory capture of vision: Examining temporal ventriloquism. *Cognitive Brain Research*, *17*(1), 154–163. https://doi.org/10.1016/S0926-6410(03)00089-2

Munoz, N. E., & Blumstein, D. T. (2012). Multisensory perception in uncertain environments. *Behavioral Ecology*, *23*(3), 457–462. https://doi.org/10.1093/beheco/arr220

Nakao, T., Osumi, T., Ohira, H., Kasuya, Y., Shinoda, J., Yamada, J., & Northoff, G. (2010). Medial prefrontal cortex-dorsal anterior cingulate cortex connectivity during behavior selection without an objective correct answer. *Neuroscience Letters*, *482*(3), 220–224. https://doi.org/10.1016/j.neulet.2010.07.041

Nava, E., Gamberini, C., Berardis, A., & Bolognini, N. (2018). Action shapes the sense of body ownership across human development. *Frontiers in Psychology*, *0*(DEC), 2507. https://doi.org/10.3389/FPSYG.2018.02507

Newen, A., & Vetter, P. (2017). Why cognitive penetration of our perceptual experience is still the most plausible account. *Consciousness and Cognition*, *47*, 26–37. https://doi.org/10.1016/j.concog.2016.09.005

Noel, J. P., Blanke, O., & Serino, A. (2018). From multisensory integration in peripersonal space to bodily self-consciousness: From statistical regularities to statistical inference. *Annals of the New York Academy of Sciences*, *1426*(1), 146–165. https://doi.org/10.1111/nyas.13867

Noel, J. P., Pfeiffer, C., Blanke, O., & Serino, A. (2015). Peripersonal space as the space of the bodily self. *Cognition*, *144*, 49–57. https://doi.org/10.1016/j.cognition.2015.07.012

Northoff, G. (2016). Is the self a higher-order or fundamental function of the brain? The “basis model of self-specificity” and its encoding by the brain’s spontaneous activity. *Cognitive Neuroscience*, *7*(1–4), 203–222. https://doi.org/10.1080/17588928.2015.1111868

Northoff, G., & Hayes, D. J. (2011). Is our self nothing but reward? *Biological Psychiatry*, *69*(11), 1019–1025. https://doi.org/10.1016/j.biopsych.2010.12.014

Northoff, G., & Panksepp, J. (2008). The trans-species concept of self and the subcortical-cortical midline system. *Trends in Cognitive Sciences*, *12*(7), 259–264. https://doi.org/10.1016/j.tics.2008.04.007

Nuerk, H. C., Wood, G., & Willmes, K. (2005). The universal SNARC effect: The association between number magnitude and space is amodal. *Experimental Psychology*, *52*(3), 187–194. https://doi.org/10.1027/1618-3169.52.3.187

O' Dowd, A., Sorgini, F., & Newell, F. N. (2020). Seeing an image of the hand affects performance on a crossmodal congruency task for sequences of events. *Consciousness and Cognition*, *80*, 102900. https://doi.org/10.1016/j.concog.2020.102900

O'Dowd, A., & Newell, F. N. (2020). The rubber hand illusion is influenced by self-recognition. *Neuroscience Letters*, *720*. https://doi.org/10.1016/j.neulet.2020.134756

Odegaard, B., & Shams, L. (2016). The brain's tendency to bind audiovisual signals is stable but not general. *Psychological Science*, *27*(4), 583–591. https://doi.org/10.1177/0956797616628860

Otten, M., Seth, A. K., & Pinto, Y. (2017). A social Bayesian brain: How social knowledge can shape visual perception. *Brain and Cognition*, *112*, 69–77. https://doi.org/10.1016/J.BANDC.2016.05.002

Panksepp, J., & Northoff, G. (2009). The trans-species core SELF: The emergence of active cultural and neuro-ecological agents through self-related processing within subcortical-cortical midline networks. *Consciousness and Cognition*, *18*(1), 193–215. https://doi.org/10.1016/j.concog.2008.03.002

Payne, B., Lavan, N., Knight, S., & McGettigan, C. (2020). Perceptual prioritization of self-associated voices. *British Journal of Psychology*, bjop.12479. https://doi.org/10.1111/bjop.12479

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

Pellencin, E., Paladino, M. P., Herbelin, B., & Serino, A. (2018). Social perception of others shapes one's own multisensory peripersonal space. *Cortex*, *104*, 163–179. https://doi.org/10.1016/j.cortex.2017.08.033

Piazza, M., Mechelli, A., Price, C. J., & Butterworth, B. (2006). Exact and approximate judgements of visual and auditory numerosity: An fMRI study. *Brain Research*, *1106*(1), 177–188. https://doi.org/10.1016/j.brainres.2006.05.104

Picard, F., & Friston, K. (2014). Predictions, perception, and a sense of self. *Neurology*, *83*(12), 1112–1118. https://doi.org/10.1212/WNL.0000000000000798

Prinzmetal, W., McCool, C., & Park, S. (2005). Attention: Reaction time and accuracy reveal different mechanisms. *Journal of Experimental Psychology: General*, *134*(1), 73–92. https://doi.org/10.1037/0096-3445.134.1.73

Prinzmetal, W., Zvinyatskovskiy, A., Gutierrez, P., & Dilem, L. (2009). Voluntary and involuntary attention have different consequences: The effect of perceptual difficulty. *Quarterly Journal of Experimental Psychology*, *62*(2), 352–369. https://doi.org/10.1080/17470210801954892

Pyasik, M., Tieri, G., & Pia, L. (2020). Visual appearance of the virtual hand affects embodiment in the virtual hand illusion. *Scientific Reports 2020 10:1*, *10*(1), 1–11. https://doi.org/10.1038/s41598-020-62394-0

Recanzone, G. H. (2003). Auditory influences on visual temporal rate perception. *Journal of Neurophysiology*, *89*(2), 1078–1093. https://doi.org/10.1152/jn.00706.2002

Reuther, J., & Chakravarthi, R. (2017). Does self-prioritization affect perceptual processes? *Visual Cognition*, *25*(1–3), 381–398. https://doi.org/10.1080/13506285.2017.1323813

Rohe, T., & Noppeney, U. (2018). Reliability-weighted integration of audiovisual signals can be modulated by top-down attention. *ENeuro*, *5*(1), ENEURO.0315-17.2018. https://doi.org/10.1523/ENEURO.0315-17.2018

Samad, M., Chung, A. J., & Shams, L. (2015). Perception of body ownership is driven by Bayesian sensory inference. *PLoS ONE*, *10*(2), e0117178. https://doi.org/10.1371/journal.pone.0117178

Schäfer, S., Frings, C., & Wentura, D. (2016). About the composition of self-relevance: Conjunctions not features are bound to the self. *Psychonomic Bulletin and Review*, *23*(3), 887–892. https://doi.org/10.3758/s13423-015-0953-x

Schäfer, S., Wesslein, A. K., Spence, C., & Frings, C. (2020). When self-prioritization crosses the senses: Crossmodal self-prioritization demonstrated between vision and touch. *British Journal of Psychology*, bjop.12483. https://doi.org/10.1111/bjop.12483

Schäfer, S., Wesslein, A. K., Spence, C., Wentura, D., & Frings, C. (2016). Self-prioritization in vision, audition, and touch. *Experimental Brain Research*, *234*(8), 2141–2150. https://doi.org/10.1007/s00221-016-4616-6

Scheller, M., Petrini, K., & Proulx, M. J. (2018). Perception and interactive technology. *Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience*, 1–50. https://doi.org/10.1002/9781119170174.epcn215

Scheller & Sui (2021, March 3). Effects of self-relevance on cross-modal cue integration. Retrieved from osf.io/rqkht

Scheller, M., & Sui, J. (2022). The power of the self: Anchoring information processing across contexts. *Journal of Experimental Psychology: Human Perception and Performance,* Forthcoming.

Serino, A. (2019). Peripersonal space (PPS) as a multisensory interface between the individual and the environment, defining the space of the self. *Neuroscience and Biobehavioral Reviews*, *99*, 138–159. https://doi.org/10.1016/j.neubiorev.2019.01.016

Shams, L., Kamitani, Y., & Shimojo, S. (2000). Illusions: What you see is what you hear. *Nature*, *408*(6814), 788. https://doi.org/10.1038/35048669

Shams, L., Kamitani, Y., & Shimojo, S. (2002). Visual illusion induced by sound. *Cognitive Brain Research*, *14*(1), 147–152. https://doi.org/10.1016/S0926-6410(02)00069-1

Shams, L., & Kim, R. (2010). Crossmodal influences on visual perception. *Physics of Life Reviews*, *7*(3), 269–284. https://doi.org/10.1016/j.plrev.2010.04.006

Shams, L., Ma, W. J., & Beierholm, U. (2005). Sound-induced flash illusion as an optimal percept. *NeuroReport*, *16*(17), 1923–1927. https://doi.org/10.1097/01.wnr.0000187634.68504.bb

Shimojo, S., Scheier, C., Nijhawan, R., Shams, L., Kamitani, Y., & Watanabe, K. (2001). Beyond perceptual modality: Auditory effects on visual perception. *Acoustical Science and Technology*, *22*(2), 61–67. https://doi.org/10.1250/ast.22.61

Siebold, A., Weaver, M. D., Donk, M., & van Zoest, W. (2015). Social salience does not transfer to oculomotor visual search. *Visual Cognition*, *23*(8), 989–1019. https://doi.org/10.1080/13506285.2015.1121946

Spagna, A., Mackie, M. A., & Fan, J. (2015). Supramodal executive control of attention. *Frontiers in Psychology*, *6*, 65. https://doi.org/10.3389/fpsyg.2015.00065

Spence, C., & Driver, J. (1996). Audiovisual Links in Endogenous Covert Spatial Attention. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(4), 1005–1030. https://doi.org/10.1037/0096-1523.22.4.1005

Stein, B. E. (Ed.). (2012). *The New Handbook of Multisensory Processing*. MIT Press.

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

Stolte, M., Humphreys, G., Yankouskaya, A., & Sui, J. (2017). Dissociating biases towards the self and positive emotion. *Quarterly Journal of Experimental Psychology*, *70*(6), 1011–1022. https://doi.org/10.1080/17470218.2015.1101477

Störmer, V. S., McDonald, J. J., & Hillyard, S. A. (2009). Cross-modal cueing of attention alters appearance and early cortical processing of visual stimuli. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(52), 22456–22461. https://doi.org/10.1073/pnas.0907573106

Sui, J., & Han, S. (2007). Self-construal priming modulates neural substrates of self-awareness. *Psychological Science*, *18*(10), 861–866. https://doi.org/10.1111/j.1467-9280.2007.01992.x

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., & Humphreys, G. W. (2013). Self-referential processing is distinct from semantic elaboration: Evidence from long-term memory effects in a patient with amnesia and semantic impairments. *Neuropsychologia*, *51*(13), 2663–2673. https://doi.org/10.1016/j.neuropsychologia.2013.07.025

Sui, J., & Humphreys, G. W. (2015a). More of me! Distinguishing self and reward bias using redundancy gains. *Attention, Perception, and Psychophysics*, *77*(8), 2549–2561. https://doi.org/10.3758/s13414-015-0970-x

Sui, J., & Humphreys, G. W. (2015b). The Integrative Self: How Self-Reference Integrates Perception and Memory. *Trends in Cognitive Sciences*, *19*(12), 719–728. https://doi.org/10.1016/j.tics.2015.08.015

Sui, J., & Humphreys, G. W. (2017a). 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., & Humphreys, G. W. (2017b). Aging enhances cognitive biases to friends but not the self. *Psychonomic Bulletin and Review*, *24*(6), 2021–2030. https://doi.org/10.3758/S13423-017-1264-1

Sui, J., & Rotshtein, P. (2019). Self-prioritization and the attentional systems. *Current Opinion in Psychology*, *29*, 148–152. https://doi.org/10.1016/J.COPSYC.2019.02.010

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

Tacikowski, P., & Ehrsson, H. H. (2016). Preferential processing of self-relevant stimuli occurs mainly at the perceptual and conscious stages of information processing. *Consciousness and Cognition*, *41*, 139–149. https://doi.org/10.1016/j.concog.2016.02.013

Tajadura-Jiménez, A., Longo, M. R., Coleman, R., & Tsakiris, M. (2012). The person in the mirror: Using the enfacement illusion to investigate the experiential structure of self-identification. *Consciousness and Cognition*, *21*(4), 1725–1738. https://doi.org/10.1016/J.CONCOG.2012.10.004

Talsma, D., Doty, T. J., & Woldorff, M. G. (2007). Selective attention and audiovisual integration: Is attending to both modalities a prerequisite for early integration? *Cerebral Cortex*, *17*(3), 679–690. https://doi.org/10.1093/cercor/bhk016

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

Talsma, D., & Woldorff, M. G. (2005). Selective attention and multisensory integration: Multiple phases of effects on the evoked brain activity. *Journal of Cognitive Neuroscience*, *17*(7), 1098–1114. https://doi.org/10.1162/0898929054475172

Tang, X., Wu, J., & Shen, Y. (2016). The interactions of multisensory integration with endogenous and exogenous attention. *Neuroscience and Biobehavioral Reviews*, *61*, 208–224. https://doi.org/10.1016/j.neubiorev.2015.11.002

Teneggi, C., Canzoneri, E., Di Pellegrino, G., & Serino, A. (2013). Social Modulation of Peripersonal Space Boundaries. *Current Biology*, *23*(5), 406–411. https://doi.org/10.1016/J.CUB.2013.01.043

Ting, C. C., Wu, S. W., Wu, S. W., Yu, C. C., Maloney, L. T., Maloney, L. T., & Maloney, L. T. (2015). Neural Mechanisms for Integrating Prior Knowledge and Likelihood in Value-Based Probabilistic Inference. *Journal of Neuroscience*, *35*(4), 1792–1805. https://doi.org/10.1523/JNEUROSCI.3161-14.2015

Truong, G., Roberts, K. H., & Todd, R. M. (2017). I saw mine first: A prior-entry effect for newly acquired ownership. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(1), 192–205. https://doi.org/10.1037/xhp0000295

Tsakiris, M. (2008). Looking for Myself: Current Multisensory Input Alters Self-Face Recognition. *PLOS ONE*, *3*(12), e4040. https://doi.org/10.1371/JOURNAL.PONE.0004040

Tsakiris, M. (2010). My body in the brain: A neurocognitive model of body-ownership. *Neuropsychologia*, *48*(3), 703–712. https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2009.09.034

Tsakiris, M. (2017). The multisensory basis of the self: From body to identity to others: *Https://Doi.Org/10.1080/17470218.2016.1181768*, *70*(4), 597–609. https://doi.org/10.1080/17470218.2016.1181768

Verplanken, B., & Sui, J. (2019). Habit and identity: Behavioral, cognitive, affective, and motivational facets of an integrated self. *Frontiers in Psychology*, *10*, 1504. https://doi.org/10.3389/fpsyg.2019.01504

Vossel, S., Geng, J. J., & Fink, G. R. (2014). Dorsal and ventral attention systems: Distinct neural circuits but collaborative roles. *Neuroscientist*, *20*(2), 150–159. https://doi.org/10.1177/1073858413494269

Wahn, B., & König, P. (2017). Is attentional resource allocation across sensory modalities task-dependent? *Advances in Cognitive Psychology*, *13*(1), 83–96. https://doi.org/10.5709/acp-0209-2

Wahn, B., Rohe, T., Gearhart, A., Kingstone, A., & Sinnett, S. (2020). Performing a task jointly enhances the sound-induced flash illusion. *Quarterly Journal of Experimental Psychology*, *73*(12), 2260–2271. https://doi.org/10.1177/1747021820942687

Walsh, V. (2003). A theory of magnitude: Common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences*, *7*(11), 483–488. https://doi.org/10.1016/j.tics.2003.09.002

Wang, A., Sang, H., He, J., Sava-Segal, C., Tang, X., & Zhang, M. (2019). Effects of Cognitive Expectation on Sound-Induced Flash Illusion. *Perception*, *48*(12), 1214–1234. https://doi.org/10.1177/0301006619885796

Witt, J. K., Taylor, J. E. T., Sugovic, M., & Wixted, J. T. (2016). Further clarifying signal detection theoretic interpretations of the Müller-Lyer and sound-induced flash illusions. *Journal of Vision*, *16*(11), 19–19. https://doi.org/10.1167/16.11.19

Woźniak, M., & Hohwy, J. (2020). Stranger to my face: Top-down and bottom-up effects underlying prioritization of images of one’s face. *PLOS ONE*, *15*(7), e0235627. https://doi.org/10.1371/journal.pone.0235627

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

Wutz, A., Melcher, D., & Samaha, J. (2018). Frequency modulation of neural oscillations according to visual task demands. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(6), 1346–1351.

https://doi.org/10.1073/pnas.1713318115

Social relevance modulates multisensory integration

- Supplementary material -

Meike Scheller[1,2], Jie Sui[1]

[1]School of Psychology, University of Aberdeen

[2]Department of Psychology, Durham University

The study has been pre-registered on the Open Science Framework (OSF). Registration, stimuli and further materials necessary to run the experiment, as well as data and analysis scripts are available through the OSF project page: https://osf.io/z2gtm/

Supplementary material

S.1: Data analysis

The three main research questions and hypotheses outlined in 1.6 of the manuscript mapped onto the three experiments in order to maintain sufficient power for the investigation of social-association effects on multisensory processing. Please refer to Table S1 or section 1.6 (main manuscript) for details.

Table S1: Outline of experiments testing hypotheses 1-3

Experiment number	**Decision criterion**	**sample *n***	**Hypotheses (sensory conditions)**
Experiment 1	Social (self/other)	40	H1 (Audio-visual congruent) H2 (Auditory only, Visual only)
Experiment 2 (Replication)	Social (self/other)	40	H1 (Audio-visual congruent) H2 (Auditory only, Visual only)
Experiments 1 & 2	Social (self/other)	80	H3 (Audio-visual incongruent)
Exeriment 3 (Control)	Object (fork/spoon)	40	H1 (Audio-visual congruent) H2 (Auditory only, Visual only) (H3 (Audio-visual incongruent))

All statistical analyses were carried out in R (version 4.0.3) using the following packages: 'simR' for power estimation and sample size calculation (Green & Macleod, 2016), 'lme4' for devising mixed effect models (Bates, Mächler, et al., 2015), 'emmeans' for follow-up contrasts (Lenth, 2020), and 'ggplot2' and 'sjPlot' for visualization of mixed effect model estimates (Lüdecke, 2017; Wickham, 2016).

While the matching paradigm requires the presentation of matching and mismatching labels (to allow individuals to make a decision along the dimension of interest, i.e. number/social) the analyses focussed only on match-trials. This is in line with a number of previous studies that reported consistent effects of self-prioritization in match-trials (e.g. Sui et al., 2012; Sui & Humphreys, 2017; Verplanken & Sui, 2019) while the reported effects of self-association on mismatch trials are less consistent. The main reason for excluding mismatch trials from the present analysis is that the study investigates how self-prioritization effects influence perceptual processing, which is typically only expressed in match-trials. Analysing mismatch-trials would have also added a further layer of difficulty in that it induced conflicts at different processing levels that were not the main focus of investigation here – that is, intersecting perceptual congruency and category matching. In other words, while in matching trials performance can be directly attributed to the perceptual information presented in the task, performance in the mismatching trials might result from different mechanism (see also Enock et al., 2020). In fact, accuracy of mismatch responses might result from differences in the perceptual information (congruent or incongruent presentation) as well as their interaction of the label with either the auditory or visual information. For example, as participants are instructed to attend to the visual information, but auditory information affects their percept, an accuracy modulation as a result of bimodal incongruence is expected (Shams et al., 2002). When presenting a visual--mismatching label in an incongruent trial, the label matches the auditory information. Participants may then base their decision either on the match with the auditory information, or the mismatch with the visual information (while the measured response is identical). Furthermore, even in unisensory or bimodal congruent trials it is not possible to disentangle to what extent the decision was based on conflict with the perceptual event numerosity or the label. Disentangling these options was not the main focus in the present study. On match-trials, the label was always in line with the information that participants are required to report, and therefore offered a way to assess perceptual changes (via congruency modulation) directly.

As outlined in our preregistration, and to avoid overparameterization while maximizing comparability between the sensory conditions, the most parsimonious random effects structure was selected for our generalized linear mixed effects models (Bates, Kliegl, et al., 2015), allowing the intercept to vary for each individual participant. The same random effect structure was used for effect size estimation based on which power was simulated (see Participants section). However, in order to explore whether a more complex random effects structure may be more suitable to describe the data, we specified models that included stimulus number as a random slope, which was allowed to vary within each individual participant. Notably, depending on individual hardware configuration and perceptual differences between participants, it is likely that two and three events are not perceived similarly well across participants. Indeed, including stimulus number as a random slope enhanced model fit in all models ($p < .03$) except in one condition (auditory condition in Experiment 1 ; $p = .141$). Furthermore, task type was added as a random slope, however, this led to overparametrization of three models and only improved fit in three other models. To allow for more direct comparability between sensory conditions and select the most parsimonious structure, the final, more complex random effects structure comprised stimulus number (slope) and subject ID (intercept). Table S.1.1 shows the summary statistic for those models, showing that, despite accounting for within-participant variance in numerosity perception, the results of the main manuscript do not change. Note that fission and fusion trials do not benefit from random slopes across stimulus numbers as the split of the incongruent condition into two and three flashes results in a between-participant comparison of identity under a constant number.

Table S1.1: Summary statistics of the interaction effects from generalized linear mixed effect models with stimulus number added as a random slope and participant ID added as a random intercept. Asterisks indicate statistically significant effects.

Experiment	Sensory condition	*OR*	*CI*	*p*
Exp 1	Bimodal congruent	2.103	1.309 – 3.380	0.002*
	Auditory	2.879	1.294 – 6.405	0.010*
	Visual	1.183	0.659 – 2.126	0.573
Exp 2	Bimodal congruent	2.134	1.386 – 3.284	0.001*
	Auditory	2.694	1.272 – 5.706	0.010*
	Visual	1.243	0.741 – 2.088	0.410
Exp 3	Bimodal congruent	0.828	0.570 – 1.204	0.323
	Auditory	0.748	0.416 – 1.343	0.331
	Visual	0.828	0.493 – 1.392	0.477

S.2: Results

S.2.1 Main result tables

Tables S2.1.1: Model result tables for social manipulation and task type in experiments 1 and 2

Experiment 1

Bimodal congruent

	Correct		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	13.03	9.15 – 18.53	**<0.001**
Identity [self]	1.778	1.252 – 2.525	**0.001**
Tasktype [social]	0.489	0.370 – 0.647	**<0.001**
Identity [self] * Tasktype	2.031	1.261 – 3.272	**0.004**
Random Effects			
σ^2	3.29		
$\tau_{00\ ID}$	0.73		
ICC	0.18		
N_{ID}	40		
Observations	3989		
Marginal R^2 / Conditional R^2	0.065 / 0.234		

Visual only

	Correct		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	5.03	3.58 – 7.05	**<0.001**
Identity [self]	2.063	1.360 – 3.128	**0.001**
Tasktype [social]	0.792	0.555 – 1.129	0.198
Identity [self] * Tasktype	1.156	0.649 – 2.058	0.622
Random Effects			
σ^2	3.29		
$\tau_{00\ ID}$	0.45		
ICC	0.12		
N_{ID}	40		
Observations	1595		
Marginal R^2 / Conditional R^2	0.043 / 0.158		

Auditory only

	Correct		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	20.73	11.99 – 35.83	**<0.001**
Identity [self]	1.102	0.631 – 1.922	0.734
Tasktype [social]	0.45	0.276 – 0.735	**0.001**
Identity [self] * Tasktype	2.83	1.309 – 6.118	**0.008**
Random Effects			
σ^2	3.29		
$\tau_{00\ ID}$	1.16		
ICC	0.26		
N_{ID}	40		
Observations	1595		
Marginal R^2 / Conditional R^2	0.039 / 0.289		

Experiment 2

Bimodal congruent

	Correct		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	10.54	8.06 – 13.78	**<0.001**
Identity [self]	0.912	0.679 – 1.225	0.539
Tasktype [social]	0.761	0.572 – 1.014	0.062
Identity [self] * Tasktype	2.114	1.374 – 3.253	**0.001**
Random Effects			
σ^2	3.29		
$\tau_{00\ ID}$	0.25		
ICC	0.07		
N_{ID}	40		
Observations	3991		
Marginal R^2 / Conditional R^2	0.016 / 0.087		

Visual only

	Correct		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	4.26	3.22 – 5.64	**<0.001**
Identity [self]	1.082	0.758 – 1.543	0.666
Tasktype [social]	0.991	0.698 – 1.409	0.962
Identity [self] * Tasktype	1.235	0.741 – 2.060	0.418
Random Effects			
σ^2	3.29		
$\tau_{00\ ID}$	0.16		
ICC	0.05		
N_{ID}	40		
Observations	1597		
Marginal R^2 / Conditional R^2	0.004 / 0.050		

Auditory only

	Correct		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	19.94	12.31 – 32.32	**<0.001**
Identity [self]	0.669	0.386 – 1.162	0.154
Tasktype [SOCIAL]	0.395	0.236 – 0.661	**<0.001**
Identity [self] * Tasktype	2.664	1.288 – 5.510	**0.008**
Random Effects			
σ^2	3.29		
$\tau_{00\ ID}$	0.44		
ICC	0.12		
N_{ID}	40		
Observations	1596		
Marginal R^2 / Conditional R^2	0.029 / 0.143		

Experiment 1 & 2

Bimodal incongruent

Fission illusion (23)

	Correct		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	0.97	0.61 – 1.55	0.902
Identity [self]	1.594	0.819 – 3.102	0.17
Tasktype [social]	1.085	0.881 – 1.337	0.441
Identity [self] * Tasktype	1.045	0.777 – 1.405	0.771
Random Effects			
σ^2	3.29		
$\tau_{00\ ID}$	2.06		
ICC	0.38		
N_{ID}	80		
Observations	3983		
Marginal R^2 / Conditional R^2	0.012 / 0.392		

Fusion illusion (32)

	Correct		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	0.5	0.31 – 0.79	**0.003**
Identity [self]	1.082	0.563 – 2.079	0.813
Tasktype [social]	1.272	1.028 – 1.574	**0.027**
Identity [self] * Tasktype	1.471	1.093 – 1.979	**0.011**
Random Effects			
σ^2	3.29		
$\tau_{00\ ID}$	1.96		
ICC	0.37		
N_{ID}	80		
Observations	3978		
Marginal R^2 / Conditional R^2	0.014 / 0.382		

Tables S2.1.2: Model result tables for object manipulation and task type in experiment 3

Experiment 3

Bimodal congruent

	Correct		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	6.702	5.152 – 8.718	**<0.001**
Identity [self]	1.133	0.873 – 1.471	0.347
Tasktype [social]	1.079	0.832 – 1.399	0.565
Identity [self] * Tasktype	0.823	0.570 – 1.189	0.3
Random Effects			
σ^2	3.29		
$\tau_{00\ ID}$	0.36		
ICC	0.1		
N_{ID}	40		
Observations	3947		
Marginal R^2 / Conditional R^2	0.001 / 0.099		

Visual only

	Correct		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	4.227	3.199 – 5.585	**<0.001**
Identity [self]	1.184	0.825 – 1.699	0.36
Tasktype [social]	0.907	0.640 – 1.285	0.584
Identity [self] * Tasktype	0.837	0.508 – 1.379	0.485
Random Effects			
σ^2	3.29		
$\tau_{00\ ID}$	0.15		
ICC	0.04		
N_{ID}	40		
Observations	1582		
Marginal R^2 / Conditional R^2	0.004 / 0.046		

Auditory only

	Correct		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	7.637	4.771 – 12.225	**<0.001**
Identity [self]	0.997	0.661 – 1.502	0.987
Tasktype [social]	1.188	0.782 – 1.804	0.419
Identity [self] * Tasktype	0.747	0.417 – 1.339	0.328
Random Effects			
σ^2	3.29		
$\tau_{00\ ID}$	1.3		
ICC	0.28		
N_{ID}	40		
Observations	1565		
Marginal R^2 / Conditional R^2	0.002 / 0.285		

Experiment 3

Bimodal incongruent

Fission illusion (23)

	Correct		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	0.938	0.543 – 1.621	0.819
Identity [self]	2.563	1.201 – 5.469	**0.015**
Tasktype [social]	1.309	0.979 – 1.752	0.07
Identity [self] * Tasktype	0.95	0.629 – 1.435	0.808
Random Effects			
σ^2	3.29		
$\tau_{00\ ID}$	1.26		
ICC	0.28		
N_{ID}	40		
Observations	1967		
Marginal R^2 / Conditional R^2	0.047 / 0.311		

Fusion illusion (32)

	Correct		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	0.86	0.512 – 1.446	0.57
Identity [self]	0.679	0.320 – 1.445	0.315
Tasktype [social]	1.442	1.093 – 1.902	**0.01**
Identity [self] * Tasktype	0.941	0.630 – 1.406	0.765
Random Effects			
σ^2	3.29		
$\tau_{00\ ID}$	1.26		
ICC	0.28		
N_{ID}	40		
Observations	1966		
Marginal R^2 / Conditional R^2	0.016 / 0.288		

Tables S2.1.3: Model result tables for social manipulation and task type in experiments 1 and 2, based only on participants that showed a self-relevant performance enhancement in the bimodal congruent condition

Experiment 1 & 2

Bimodal incongruent

Fission illusion (23)

	Correct		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	1.159	0.688 – 1.954	0.579
Identity [self]	1.581	0.720 – 3.473	0.254
Tasktype [social]	0.891	0.700 – 1.135	0.351
Identity [self] * Tasktype	1.49	1.026 – 2.164	**0.036**
Random Effects			
σ^2	3.29		
$\tau_{00\ ID}$	1.82		
ICC	0.36		
N_{ID}	52		
Observations	2584		
Marginal R^2 / Conditional R^2	0.022 / 0.371		

Fusion illusion (32)

	Correct		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	0.698	0.402 – 1.213	0.202
Identity [self]	1.038	0.497 – 2.168	0.92
Tasktype [social]	1.037	0.784 – 1.373	0.798
Identity [self] * Tasktype	1.778	1.234 – 2.562	**0.002**
Random Effects			
σ^2	3.29		
$\tau_{00\ ID}$	1.58		
ICC	0.32		
N_{ID}	52		
Observations	2580		
Marginal R^2 / Conditional R^2	0.016 / 0.335		

S.2.2 Intra-individual correlation of identity- and object-biases

While, on average, groups of individuals prioritize self-related information, there is individual variation in the extent to which self- or other-related information is processed (Scheller & Sui, 2022; Stolte et al., 2017). Given that individuals may express social prioritization to different extents, we assessed whether the individual direction and magnitude of bias, which was expressed in response enhancement towards the self- or the other-associated number of flashes in the bimodal congruent condition, correlated with the degree of bias when audio-visual information was in conflict. That is, for each individual participant we calculated an accuracy change score from baseline (social condition – number condition), which was subsequently used to create a difference score between the self- and other-associated number of flashes. Positive scores indicate a bias towards self-related information, while negative scores indicate a bias towards other-associated information, as a consequence of social association. In Experiment 3, the social condition was replaced with the object condition, and positive scores indicate a bias towards fork-associated information, while negative scores indicate a bias towards spoon-associated information, as a consequence of object association.

In order to assess whether the social- or object- bias that was measured in the bimodal congruent condition was predictive of the degree to which audio-visual information was integrated in the bimodal incongruent condition, we assessed intra-individual correlation across the two tasks. Independently of whether the numerosity-association was a social identity or an object, the tendency to be biased towards either of the two categories affected the degree to which this numerosity dominated the percept when audio-visual information was in conflict (see Figure S2.2; social: $r(80) = 0.392$; objects: $r(40) = 0.435$). However, a shift of the average group bias was only seen for the social, but not the object task. Note that the baseline-correction limits the possible magnitude of improvement at high baseline accuracies, leading to a narrower spread in the bimodal congruent condition than in the bimodal incongruent condition in experiment 1 and 2 (bimodal congruent baseline accuracy was 92.8% and 90%, respectively). There was no clear separation in bias directionality and magnitude between groups 1 and 2 in all three experiments, supporting the assumption that identity- and object-related performance-enhancement was not due to the actual number of events with which the association was formed.

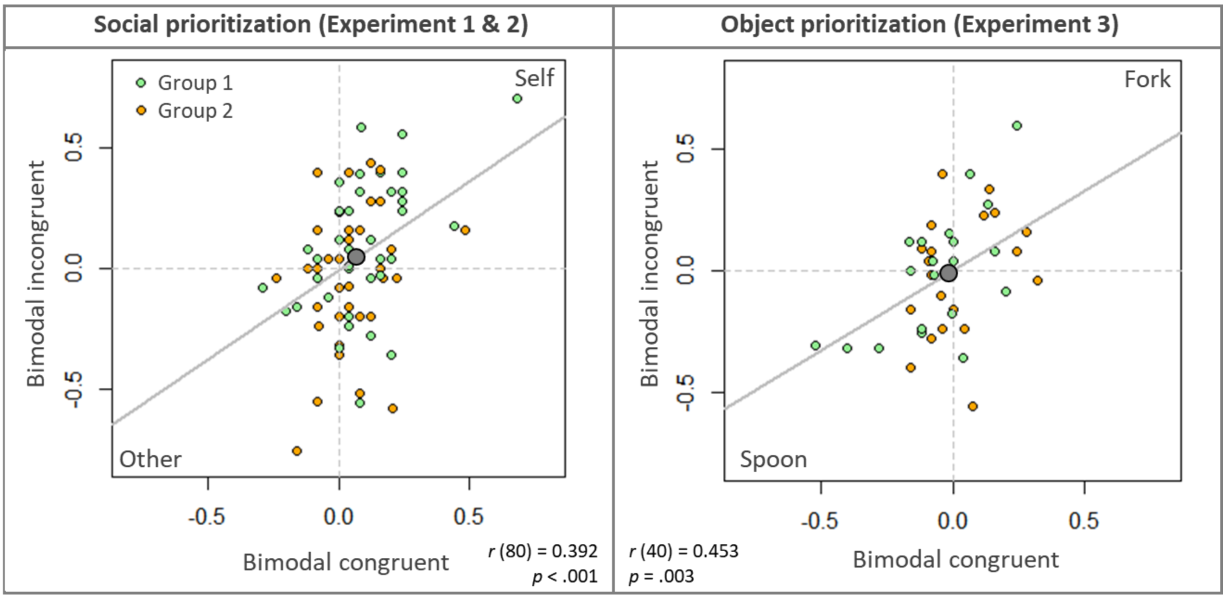

Figure S2.2: Individual accuracy-bias change scores towards either the self- or other-associated numerosity in the social condition of Experiment 1 and 2 (left), and towards the fork- or spoon-associated numerosity in the object condition of the control task (right). Individual bias scores have been baseline-corrected by subtracting individual accuracy of the number condition from the social/object condition. Group 1 and Group 2, both of which formed associations of the self/fork with two events (group 1) or three events (group 2) are color-coded in orange and green. Large dark-grey circles indicate the overall group mean, showing a deviation from the point of origin (see dashed lines) in the social prioritization task, but not in the object prioritization task. Solid grey lines indicate the best fit.

S2.3: Response time effects

Self-prioritization effects are typically expressed via an increase in accuracy as well as a speeding of response times (RTs) (Golubickis et al., 2017; Hu et al., 2020; Sui et al., 2012). The main analysis in this study was focussed on accuracy, as accurate performance is directly linked to the event numerosity percept that was subject to the social manipulation, and might therefore be more indicative of modulatory effects of endogenous attention (Prinzmetal et al., 2005, 2009). However, there is further information we can glean from assessing participants' RTs. It is important to note that RT was measured relative to the onset of the label that was either a number or a social identity/object category, in order to diminish the impact of stimulus duration differences between events of variable length. This allowed participants to form a percept of either two or three events prior to label onset. Given that the decision was based on the match or mismatch of the label and the event numerosity, it is reasonable to assume that a facilitation of self-relevant information is reflective of prediction effects, rather than purely perceptual effects. That is, as the label followed the event for which the numerosity had to be estimated, the number percept led to a prediction, which allowed for an automatic speeding of responses if the predicted (matching) label was shown (Schwager et al., 2017). Given that the analysis was focussed on match-trials, speeding of RTs towards identity-related information can be interpreted as a predictive bias toward this identity.

Data analysis

Generalized linear mixed models with inverse Gaussian were used to assess the effects in RTs. Similar to the analysis of accuracy estimates, task type (number and social) and social identity (self or other) were included as fixed factors and participant was entered as a random intercept effect. Only trials with correct responses and RTs slower than 200ms were included. Furthermore, trials that fell beyond 2 times the IQR (5.6% of trials) were excluded as outliers.

Power estimation for response time analyses

For SPE effects reflected in RT facilitation, we estimated an interaction effect of task type and associated identity of 80ms magnitude via a generalized mixed model, which was adjusted to a smaller size of 75ms for power estimation. The effect size estimate was based on pilot data from 11 naïve participants. Based on 1000 simulations, an alpha level of 0.05 and an interaction-specific effect size of 75ms, we estimated that 40 participants would allow to detect an effect of similar size with 88.6[CI95%: 86.47; 90.5]% power.

Results

Experiment 1: significant interactions of task type and social identity for the bimodal congruent ($\beta(3433) = -97.245$, $t = 10.417$, $p < .001$), visual only ($\beta(1298) = -73.281$, $t = 4.571$, $p < .001$), and auditory only ($\beta(1380) = -69.37$, $t = 4.611$, $p < .001$) conditions. See Figure S2.3.1 for visualization of marginal effects across all three experiments.

Experiment 2: significant interactions of task type and social identity for the bimodal congruent ($\beta(3489) = -35.241$, $t = 3.476$, $p < .001$) and visual only ($\beta(1262) = -43.2$, $t = 2.502$, $p = .012$) conditions, but no significant interaction in the auditory only condition ($\beta(1410) = 3.125$, $t = 0.17$ $p = .865$).

Cross-experimental analysis: For the fission illusion (two flashes, three beeps) we observed significant interactions of task type and social identity ($\beta(2151) = -98.96$, $t = 5.681$, $p < .001$). Similarly, there was a significant interaction in the fusion illusion ($\beta(1639) = -56.377$, $t = 2.97$, $p = .003$). Across both illusions, RTs were speeded towards self-associated flash numerosities compared to other-associated numerosities (see Figure S2.3.2).

Experiment 3: no interactions of task type and object for neither the bimodal congruent ($\beta(3253) = -3.171$, $t = 0.318$, $p = .751$), visual only ($\beta(1204) = 12.063$, $t = 0.759$, $p = .448$), nor auditory only conditions ($\beta(1250) = -5.577$, $t = 0.303$, $p = .762$). There were also no interactions between task type and object association either in the fission illusion (β(1152) = 5.151, t = 0.245, p = .806) nor in the fusion illusion (β(886) = 11.53, t = 0.517, p = .605).

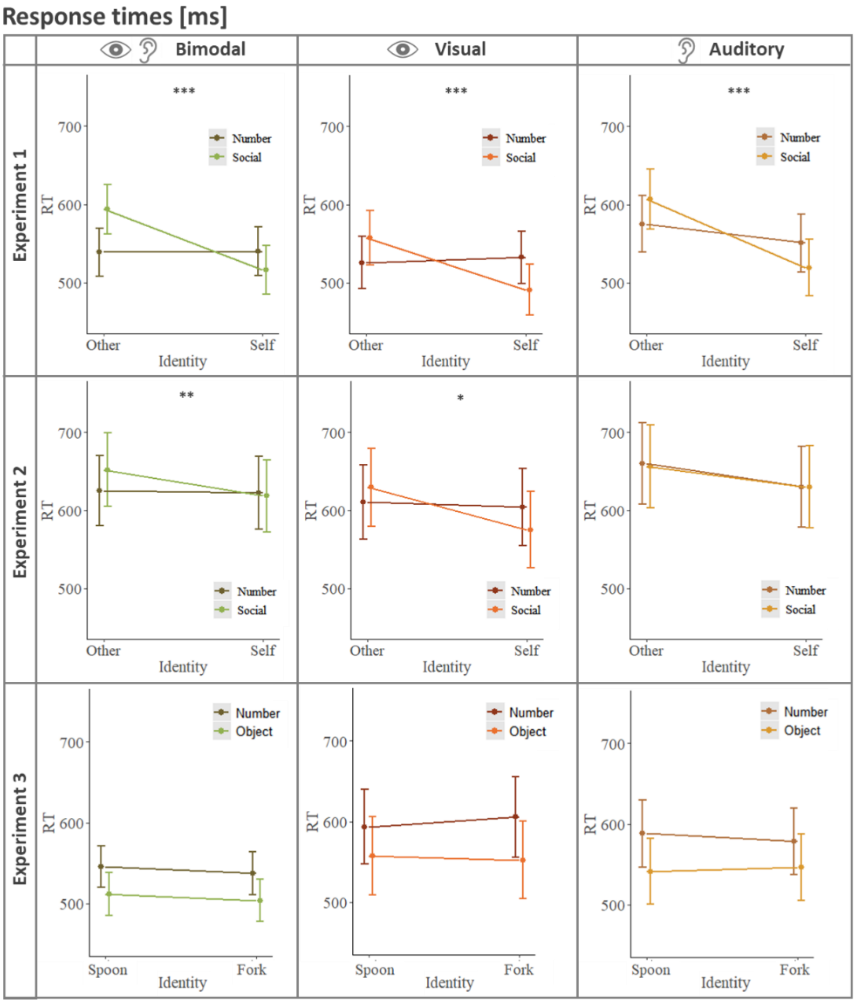

Figure S2.3.1: Marginal effects of interaction terms for RT estimates from the sequential matching tasks in the bimodal congruent and unimodal conditions. Decision judgements were made either based on the number of events (Number), or the associated social identity (Social). Error bars indicate standard error of the mean. Asterisks indicate the significance of interaction effects between social identity and task type. **p<.01; ***p<.001

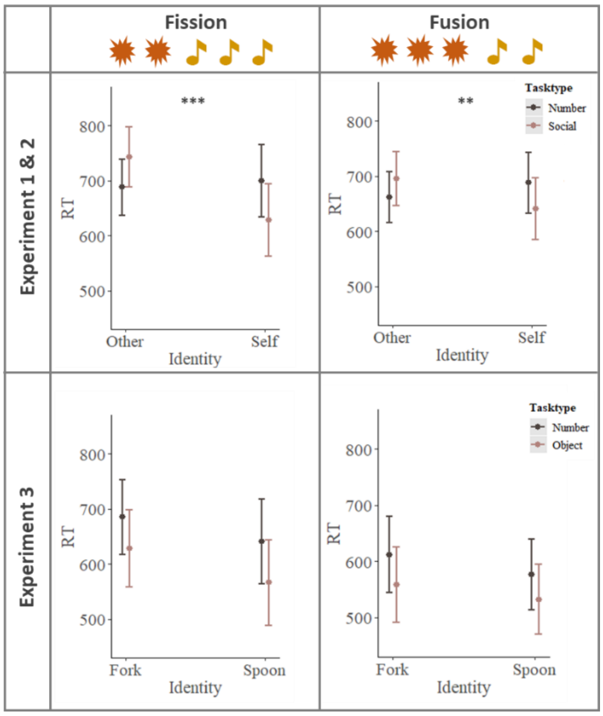

Figure S2.3.2: Marginal effects of interaction terms for RT estimates from the sequential matching tasks in the bimodal incongruent conditions. Decision judgements were made either based on the number of events (Number), or the associated social identity (Social). Error bars indicate standard error of the mean. Asterisks indicate the significance of interaction effects between social identity and task type. $^{**}p<.01$; $^{***}p<.001$

Given that the order of conditions (number first, social second) was constant, one may assume that task learning leads to a speeding of RTs in the social condition, explaining the RT facilitation of self-related information. Notably, there are several reasons to assume that this is not the case. Firstly, participants did only receive feedback during the practice trials before each task type, but not during the main task. Secondly, a task general practice effect would likely affect both social or object categories in a similar way, which is not what we observed. Instead, only the self-related number of events benefitted from faster processing. However, to ascertain that the observed effects were not affected by learning we plotted the average RTs for each of the five blocks, task types and social identities/object categories (see Figure S2.3.3). These show that, while there was an overall speeding of RTs with task practice, this did not lead to an immediate advantage of numerosity estimation via social associations.

Instead, event numerosity estimation was equally fast in both number and social task type in the first blocks. The main difference in RTs emerged merely after the association with social identities, showing an immediate RT advantage of the self- compared to the other-associated event numerosity.

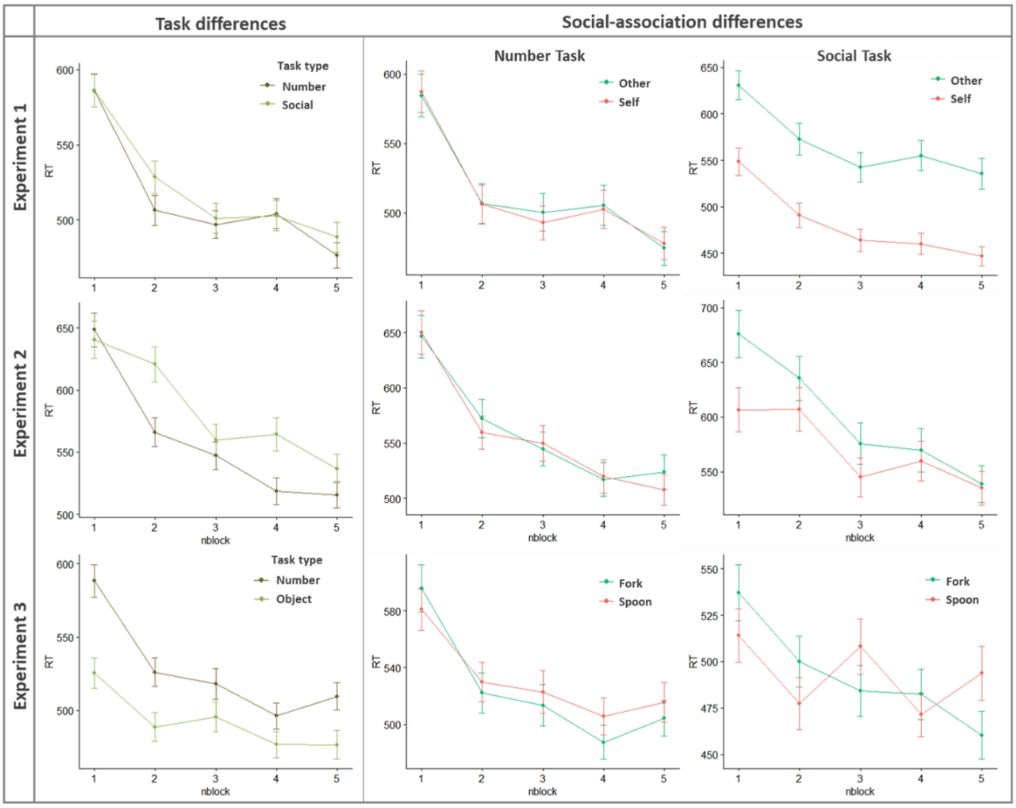

Figure S2.3.3: Mean RT (ms) changes across the five blocks in each task type (left column), and for each individual social- or object-associated category (self/other, fork/spoon) split by task type (right columns). Error bars indicate standard error of the mean.

S3: Deviations from preregistration

S3.1. Hypothesis 3: Signal detection approach – changes in perceptual sensitivity or response criterion?

In order to follow-up modulations of multisensory integration by social salience, we described a fourth hypothesis in the preregistration (H3 in main text). Here, we aimed to investigate whether socially-induced changes in illusion strength (i.e. interactions of social identity and task type) may be reflected in changes in perceptual sensitivity *d'* or bias *c*. It is important to note that, as the sound-induced flash illusion has been repeatedly shown to be perceptual in nature (for a review, see Shams & Kim, 2010), both markers are typically interpreted as perceptual changes in this illusion (Knotts & Shams, 2016; Witt et al., 2015, 2016). This hypothesis was not necessary to assess influences of social salience on the cross-modal deployment of self-relevance nor the involvement of self-relevance in multisensory integration. Instead, it offered a possible way to follow-up whether changes in the incongruent condition were reflected in the same changes in perceptual sensitivity (*d'*) or response criterion (c)that underlie the illusion. Hence, this analysis was moved to the supplementary material (see S.4 below).

S3.2. Outliers and Exclusions

As outlined in the preregistration, we specified that at participant-level, data from participants with overall accuracy-levels below 65% in the bimodal congruent conditions (for both 2 and 3 events) or in the auditory condition will be excluded from further analysis. This criterion was applied. However, we did previously not consider individual a priori biases towards specific frequencies. For example, two participants reported on over 90% of visual-only trials to hear see events, even if two events were presented. This can be explained either by perceptual bias or hardware limitations. Given the limited ability to perceive one of the two event frequencies in one of the modalities, we applied the accuracy threshold of a minimum of 65% accuracy for the visual-only trials as well, minimising the chance of including individuals with strong a priori biases.

S3.3. Intra-individual correlations

Analysing intra-individual correlations between the socially-induced biases in bimodal congruent and incongruent conditions have not been included in the preregistration. Following an exploratory sub-group analysis within the cross-experimental analysis of fission and fusion trials, it became evident that heterogeneity in SPE expression within the group might conceal any effects present across the two incongruent conditions. Indeed, individual variation in the expression of SPEs is known from the literature (Scheller & Sui, 2022; Stolte et al., 2017), and can be expected based on individual-specific differences in self-representation. However, only after evidencing that social association can modulate the integration of audio-visual cues would it make sense to explore whether this modulation may, in fact, be related to the magnitude of self-bias. Hence, while the intra-individual correlation was not necessary to show that multisensory integration is modulated by social relevance, it provides a deeper understanding of this modulation by showing that the extent to which behavioural performance is facilitated by self-association is proportional to the extent to which integration is modulated.

***S4*: Signal detection measures of discriminating flash numerosities**

Follow-up analyses probed how the social effects on the perceptual illusion, observed in our results, manifested in changes of perceptual sensitivity (*d'*) and bias (*c*). Previous studies showed that signal detection parameters vary depending on the conditions that are compared (unimodal with bimodal or different bimodal conditions; Knotts & Shams, 2016; Witt et al., 2016). While multisensory illusion susceptibility measured by contrasting unimodal and bimodal conditions typically show changes in both sensitivity (*d'*) and bias (*c*) (Knotts & Shams, 2016; Rosenthal et al., 2009; Vanes et al., 2016), perceptual biases that directly contrast the fission and fusion illusion are based on comparing bimodal conditions, and are reflected in changes of *c* (Witt et al., 2015, 2016). While contrasting uni- and bimodal conditions for every signal allows to gauge the effect of adding sounds to the specific number of flashes, indicating the cross-modal influence of sound on vision during integration (Knotts & Shams, 2016), contrasting two bimodal conditions allows to describe asymmetry in perceptual bias induced by the illusion (Witt et al., 2016). In other words, if the illusion affected the perception of a certain number of flashes more strongly than the other number of flashes, one would not only expect changes in c, but also in d'. As, in the present study, participants showed stronger susceptibility to the fusion than the fission illusion (see Table 1 in main text), such asymmetry is expected. However, if this asymmetry is further enhanced or reduced by social association is not clear.

To that end, we first calculated perceptual sensitivity and bias parameters for each unimodal (d'^{uni}_2, d'^{uni}_3, d'^{uni}_3, c^{uni}_3) and each bimodal condition (d'^{bi}_2, d'^{bi}_3, d'^{bi}_3, c^{bi}_3), and then contrasted the parameters according to the suggestions provided by Knotts and Shams (2016) and Witt et al. (2016) (see Figure S4). According to Rosenthal et al. (2009) and Knotts and Shams (2016) we would expect illusion susceptibility as a result of multisensory integration and illusion strength to show in measures *d'* and *c* when comparing unimodal and bimodal conditions. According to Witt et al. (2015, 2016), we would expect the illusion strength to be indicated by *c*, and illusion asymmetry indicated by *d'*, when comparing both multisensory conditions.

In a first instance, we established whether we could observe the predicted changes in the number (baseline) condition. Secondly, as the sound-induced flash illusion constitutes a perceptual phenomenon (Shams & Kim, 2010) we assessed how social salience modulated the parameters that were affected by the this illusion.

Perceptual sensitivity and bias were calculated based on the hits and false alarm rates for each number of flashes under a constant number of beeps. That is, the signal was either defined as 2 flashes (unimodal: 2V0A, bimodal: 2V2A) or three flashes (unimodal: 3V0A, bimodal: 3V3A). Parameter *d'2* reflects the sensitivity with which two and three flashes can be discriminated when two flashes were defined as the signal, while Parameter *d'3* reflects the sensitivity with which two and three flashes can be discriminated when three flashes were defined as the signal. Criterion *c2* and *c3* reflect the bias to report two or three flashes when either two or three flashes were defined as the signal, respectively.

Sensitivity *d'* was calculated for two and three flashes, as

$$d' = z(HR) - z(FA)$$

Hit rate (HR) and false alarm rates (FA) were defined based on a change in number of flashes under a constant number of beeps. For example, for the conditions in which 2 flashes were presented, hits were defined as correct responses in conditions where 2 flashes were presented along with 2 beeps (congruent), while false alarms were defined as the incorrect responses towards trials in which 3 flashes were presented with 2 beeps (incongruent, see Figure S4). As indicated by *z*(), hit and false alarm rates were transformed via an inverse cumulative Gaussian distribution. Cases in which individual hit or false alarm rates in a specific condition were 0 or 1 they were approximated by 1/N and 1-(1/N) respectively, with N being the number of trials tested in the respective condition.

The criterion *c* indicates the bias, i.e. the tendency to respond either “two” or “three” flashes, under the influence of a simultaneously presented number of auditory beeps (see also Keil, 2020; Knotts & Shams, 2016; Vanes et al., 2016; Witt et al., 2016), and was calculated as:

$$c = \frac{z(HR) + z(FA)}{2}$$

Positive values indicate a bias to report that the signal (i.e. two or three flashes in *c2* and *c3*, respectively) was present. As indicted by Vanes et al. (2016) and Witt et al. (2015, 2016), the influence of audition over vision in the fission and fusion illusions is likely indicated by a shift in criterion *c*. That is, the overall tendency to respond that three flashes were perceived when two flashes and three beeps were presented will enhance the fission illusion, and vice versa for the fusion illusion.

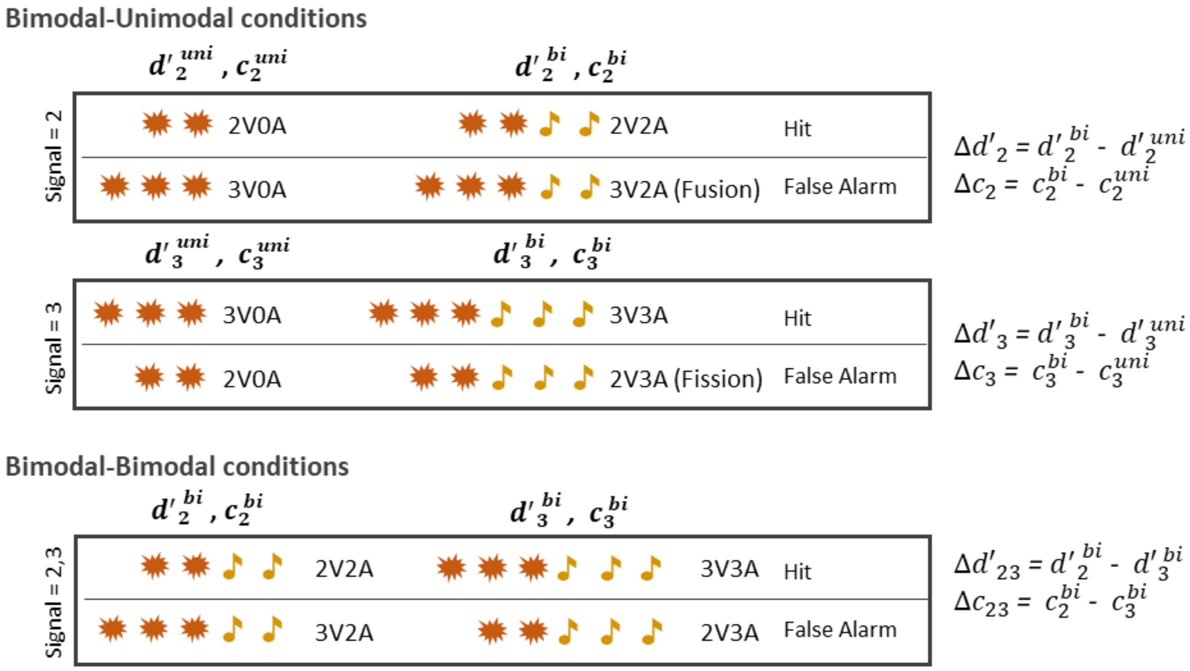

Figure S4. Specification of signal detection parameters. Differences in perceptual sensitivity *d'* and bias *c* were estimated between unimodal and bimodal conditions (Knotts & Shams, 2016; Vanes et al., 2016) and between the bimodal conditions (Witt et al., 2016). For each defined signal the hit and false alarm rates were calculated based on the outlined conditions.

S4.1. Illusion susceptibility, illusion strength and symmetry - baseline condition

Using the number (baseline) condition, we first confirmed that both participant groups did not significantly differ in their pre-existing sensitivity (0 beeps: $p = .273$; 2 beeps: $p = .812$; 3 beeps: $p = .511$) and biases (0 beeps: $p = .131$; 2 beeps: $p = .985$; 3 beeps: $p = .160$), indicating that random group-allocation was successful and allowing the grouping of individuals for this baseline analysis.

Based on Rosenthal et al. (2009) and Knotts & Shams (2016) we expect differences in sensitivity between unimodal and bimodal conditions, with smaller *d'* in bimodal compared to unimodal conditions. Indeed, we observed a reduction in sensitivity for two flashes ($\Delta d'_2$: $t(79) = -7.17$, $p < .001$) and three flashes ($\Delta d'_3$: $t(79) = -3.75$, $p < .001$) when additional sounds were presented. We also observed a change in bias from two flashes (Δc_2: $t(79) = 15.775$, $p < .001$)

and from three flashes (Δc_3: *t (79)* = 9.46, *p* < .001) towards the number of beeps when they were presented. Overall, Cohen's *d* effect sizes showed that the illusion susceptibility was more pronounced when two flashes were presented ($\Delta d'_2$: *d* = -1.011; Δc_2: *d* = 2.126) compared to when three flashes were presented ($\Delta d'_3$: *d* = -0.494; Δc_3: *d* = 1.239), evidences by a stronger decrease in sensitivity and a stronger increase in bias.

Based on Witt et al. (2015, 2016) we expected differences in bias between the two bimodal conditions (Δc_{23}) as an indicator of the sound-induced flash-illusion. Given that participants showed stronger susceptibility to the fusion than the fission illusion (see Table 1 in main text), we would predict differences in sensitivity between the two bimodal conditions ($\Delta d'_{23}$), suggesting that the illusion is asymmetrical depending on the number of flashes/beeps. Indeed, we found that both sensitivity (*t (79)* = 4.275, *p* < .001, *d* = 0.43) and the perceptual bias (*t (79)* = 3.777, *p* < .001, *d* = 0.607) significantly differed between the two illusions.

S4.1. Effects of social salience on illusion susceptibility, illusion strength and symmetry

To assess the effects of social salience on changes in illusion susceptibility, we used linear mixed effect models with participant ID as random factor to test for significant interactions of task type (number, social) and social identity association on the difference scores of unimodal and bimodal conditions ($\Delta d'_2$, $\Delta d'_3$ and Δc_2, and Δc_3) that were contrasted in the baseline task above. This analysis indicated a significant interaction in perceptual sensitivity for discriminating three flashes ($\Delta d'_3$: *β* (160) = 0.52, *p* = .012) when additional beeps were presented, but not for discriminating two flashes ($\Delta d'_2$: *β* (160) = 0.292, *p* = .147). This interaction was driven by an increase in sensitivity when three flashes were associated with the self (*β* (78) = 0.335, *p* = .025), rather than the other (*β* (78) = -0.186, *p* = .207; see Figure S4.2). Furthermore, we observed a significant interaction of task type and social association on the criterion when perceiving two flashes (Δc_2: *β* (160) = -0.21, *p* = .032), but not when perceiving three flashes (Δc_3: *β* (160) = 0.05, *p* = .663). This interaction was driven by a stronger bias when the other was associated with two flashes (*β* (78) = 0.335, *p* < .001), compared to when the self was associated with two flashes (*β* (78) = 0.04, *p* = .570; see Figure S4.2). Together this suggests that adding an auditory stimulus lead to the enhancement of perceptual sensitivity when three flashes were associated with the self, but not with the other, and that the bias induced by auditory stimuli

Lastly, we compared sensitivity and criterion between the two bimodal conditions to assess the impact of social association on the SIFI illusion strength and symmetry. To do so, we used linear mixed effect models with participant ID as random factor to test for significant interactions of task type (number, social) and social identity association on the difference scores of the two bimodal conditions that were contrasted in the baseline task above. We observed no significant interaction for sensitivity (*β* (160) = -0.23, *p* = .226) but a significant effect on the illusion bias (*β* (160) = -0.44, *p* = .001). This interaction effect was based on a stronger bias towards three flashes when three flashes were associated with the self and two flashes were associated with the other (*β* (78) = -0.39, *p* < .001) rather than the other way around (*β* (78) = -0.054, *p* < .579; see Figure S4.2).

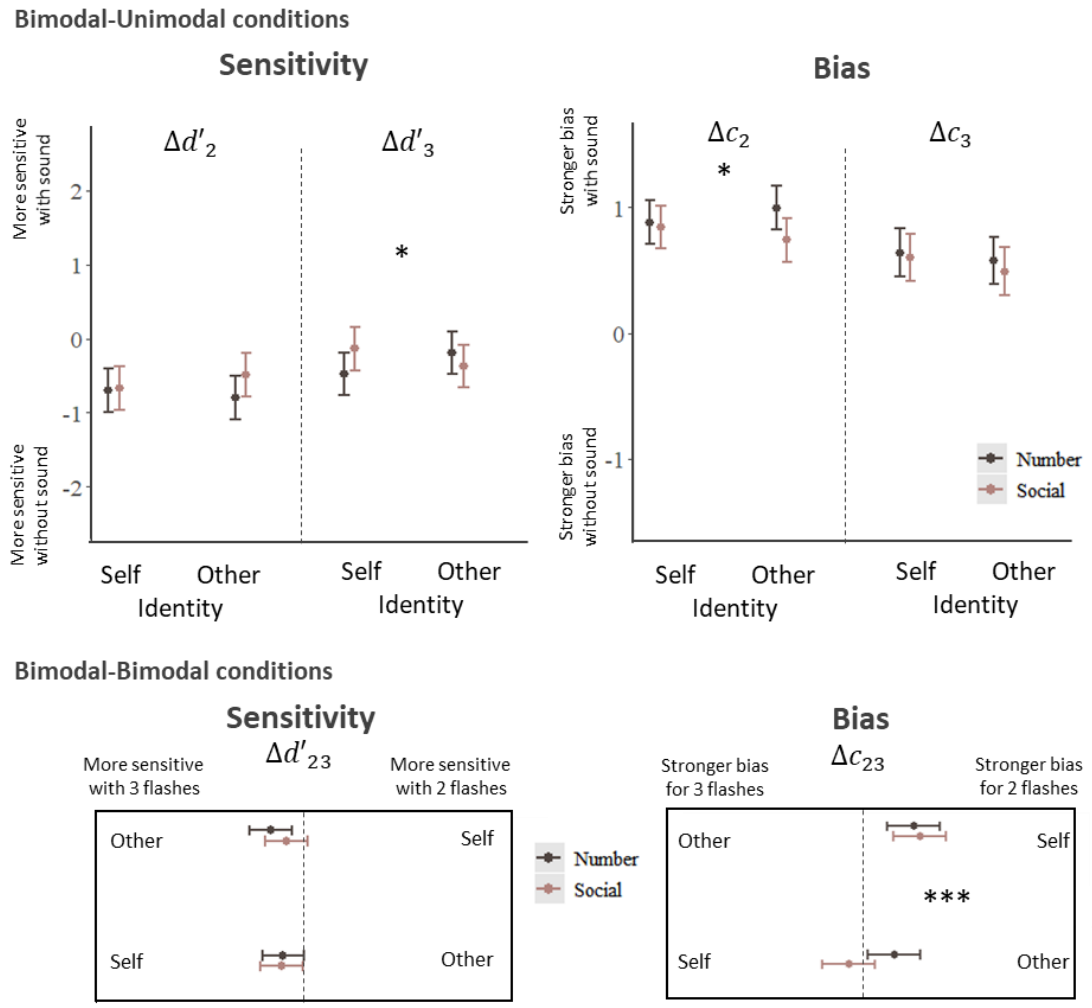

Figure S4.2: Top panel: Difference scores of perceptual sensitivity *d'* and bias *c* between bimodal and unimodal conditions for two and three flashes, depending on social association and task type. Higher *d'* indicates a higher sensitivity to respond to the respective number of flashes when auditory beeps were presented. Higher values of *c* indicate a greater tendency to respond that the signal (2 or 3 flashes) was present when auditory beeps were presented. Lower panel: Difference scores of perceptual sensitivity *d'* and bias *c* between the bimodal conditions. Dashed lines indicate equality lines. Social identities within each graph correspond to the identities associated with each respective number of flashes. Dark markers represent results from the number (baseline) task, while light markers represent results from the social task. Error bars indicate SEM. Asterisks indicate the significance of interaction effects between social identity and task type. *$p<.05$ ***$p<.001$

Taken together, our signal detection results showed that, in general, participants were less sensitive and showed a stronger bias when sounds were added. As the addition of congruent information typically leads to an enhancement of performance, while the addition of incongruent information decreases performance (see also Table 1), this difference was likely driven by the higher false alarm rate for illusion conditions. The proportion of fission illusions was lower than fusion illusions in our sample, which is in line with the higher sensitivity to three flashes than to two flashes when accompanied by auditory stimuli (baseline $\Delta d'_2$ < $\Delta d'_3$). Notably, this benefit may, in fact, result from the higher uncertainty of visually perceiving three flashes in the unimodal conditions. Given the relatively lower sensitivity in correctly detecting two flashes when presented alongside auditory beeps (baseline $\Delta d'_2$), the auditory-induced bias to report two flashes in bimodal conditions was accordingly higher than the bias to report three flashes (baseline $\Delta c_2 > \Delta c_3$), reflected in a stronger fusion illusion. However while the comparison of unimodal and bimodal conditions showed that the fusion illusion was stronger, it was also highly significant for the fission illusion (i.e. baseline $\Delta d'_2$,>0; baseline Δc_2 >0) suggesting that the added sounds led to the integration of auditory and visual information and an enhancement of illusion percepts, reflected in both a significant effect on perceptual sensitivity and bias. When contrasting the two bimodal conditions in the number (baseline) conditions, perceptual sensitivity and bias suggested an Illusion asymmetry that was also evidenced by the unimodal and bimodal contrasts: here, three flashes were processed with higher sensitivity when accompanied by sounds (baseline $\Delta d'_{23}$) and participants showed an overall stronger perceptual bias towards two flashes (baseline Δc_{23}), indicative of the fusion illusion. Taken together, these markers suggest an asymmetry in the illusion effect, i.e. that fission and fusion are not equally strong.

Notably, the sensitivity to three flashes was enhanced when three flashes were associated with the self and accompanied by beeps ($\Delta d'_3$), suggesting that participants benefitted more from the additional auditory information when it was congruent (associated with the self) and more often correctly ignored it when it was incongruent (associated with the other). At the same time, participants showed a reduced bias to report two flashes when two flashes were associated with the other (Δc_2). This was, again, likely driven by the influence of the self-associated auditory beeps in the fusion illusion, which led to an increased false alarm rate of three flashes when two flashes were coupled with three beeps. In fact, previous studies showed that self-relevant information is more difficult to ignore, even if it impairs task performance (Brédart et al., 2006; Gronau et al., 2003). Lastly, while the existing asymmetry in sensitivity towards the fission or fusion illusion was not altered by self-association ($\Delta d'_{23}$), the strength of the existing perceptual bias was reduced by self-association (Δc_{23}), but only when the self was associated with three flashes, which is in line with the enhanced sensitivity to three flashes shown in the unimodal vs bimodal contrast ($\Delta d'_3$). Taken together, the contrast of unimodal and bimodal conditions showed a pattern that would be expected if the fusion illusion was reduced via a self-association-induced enhancement of sensitivity towards three events (flashes and beeps). Notably, while the perceptual bias resulting in the fusion illusion was reduced by self-association, the existing asymmetry of the illusion was not significantly affected. This would suggest that the effect of social association did not only affect the fusion illusion, but was proportionally stronger than in the fission illusion. Exploratory follow-up analyses on a subgroup of individuals that showed a self-association benefit in the bimodal congruent conditions indeed indicated a significant effect of social association on the fission illusion (see tables S.2.1.3).

In Experiment 3 (control), there were no interaction effects of task type and object identity association on perceptual sensitivity, neither for $\Delta d'_2$ ($\beta(80) = 0.29$, $p = .290$) nor $\Delta d'_3$ ($\beta(80) = 0.32$, $p = 0.186$). Furthermore, there was no interaction effect on the perceptual bias for two flashes (Δc_2: $\beta(80) = 0.02$, $p = .893$) nor for three flashes (Δc_3: $\beta(80) < 0.01$, $p = .980$). Also, when contrasting the two bimodal conditions there was no interaction in sensitivity ($\beta(80) = -0.03$, $p = .913$) nor in bias ($\beta(80) = 0.10$, $p = .645$), supporting the main findings that the association with arbitrary household items did not systematically bias audio-visual temporal numerosity perception.

Overall, the results from the signal detection analysis suggest that the difference in expression of self-related modulations in the fusion, but not the fission illusion may result from the illusion asymmetry, which, in turn, may be determined by the sensory uncertainty of the stimulus. That is, the attention shift induced by self-association may enhance the appearance of the perceptual representation by reducing the uncertainty of information at relatively early stages of processing (Rohe & Noppeney, 2018), leading to a shift of weights during integration that is larger, i.e. more readily measurable, than the shift of weight in an already reliable signal. Future studies should target the SIFI asymmetry to elucidate why different audio-visual temporal illusions are differentially prone to attentional modulations through social salience.

References

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). *Parsimonious mixed models*. http://arxiv.org/abs/1506.04967

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

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

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

Golubickis, M., Falben, 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 and Cognition*, *45*(7), 1223–1239. https://doi.org/10.3758/s13421-017-0722-3

Green, P., & Macleod, C. J. (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, *7*(4), 493–498. https://doi.org/10.1111/2041-210X.12504

Gronau, N., Cohen, A., & Ben-Shakhar, G. (2003). Dissociations of personally significant and task-relevant distractors inside and outside the focus of attention: A combined behavioral and psychophysiological study. *Journal of Experimental Psychology: General*, *132*(4), 512–529. https://doi.org/10.1037/0096-3445.132.4.512

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

Keil, J. (2020). Double flash illusions: Current findings and future directions. *Frontiers in Neuroscience*, *14*, 298. https://doi.org/10.3389/fnins.2020.00298

Knotts, J. D., & Shams, L. (2016). Clarifying signal detection theoretic interpretations of the Müller-Lyer and sound-induced flash illusions. *Journal of Vision*, *16*(11), 18–18. https://doi.org/10.1167/16.11.18

Lenth, R. (2020). *emmeans: Estimated Marginal Means, aka Least-Squares Means.* (1.5.3). CRAN.

Lüdecke, D. (2017). sjPlot: Data visualization for statistics in social science, R package version 2.4.0. *R Package Version 2.4.0*. https://cran.r-project.org/package=sjPlot

Prinzmetal, W., McCool, C., & Park, S. (2005). Attention: Reaction time and accuracy reveal different mechanisms. *Journal of Experimental Psychology: General*, *134*(1), 73–92. https://doi.org/10.1037/0096-3445.134.1.73

Prinzmetal, W., Zvinyatskovskiy, A., Gutierrez, P., & Dilem, L. (2009). Voluntary and involuntary attention have different consequences: The effect of perceptual difficulty. *Quarterly Journal of Experimental Psychology*, *62*(2), 352–369. https://doi.org/10.1080/17470210801954892

Rohe, T., & Noppeney, U. (2018). Reliability-weighted integration of audiovisual signals can be modulated by top-down attention. *ENeuro*, *5*(1), ENEURO.0315-17.2018. https://doi.org/10.1523/ENEURO.0315-17.2018

Rosenthal, O., Shimojo, S., & Shams, L. (2009). Sound-induced flash illusion is resistant to feedback training. *Brain Topography*, *21*(3–4), 185–192. https://doi.org/10.1007/S10548-009-0090-9/FIGURES/3

Scheller, M., & Sui, J. (2022). The power of the self: Anchoring information processing across contexts. *Journal of Experimental Psychology: Human Perception and Performance,* Forthcoming.

Schwager, S., Gaschler, R., Rünger, D., & Frensch, P. A. (2017). Tied to expectations: Predicting features speeds processing even under adverse circumstances. *Memory and Cognition*, *45*(4), 611–624. https://doi.org/10.3758/s13421-016-0683-y

Shams, L., Kamitani, Y., & Shimojo, S. (2002). Visual illusion induced by sound. *Cognitive Brain Research*, *14*(1), 147–152. https://doi.org/10.1016/S0926-6410(02)00069-1

Shams, L., & Kim, R. (2010). Crossmodal influences on visual perception. *Physics of Life Reviews*, *7*(3), 269–284. https://doi.org/10.1016/j.plrev.2010.04.006

Stolte, M., Humphreys, G., Yankouskaya, A., & Sui, J. (2017). Dissociating biases towards the self and positive emotion. *Quarterly Journal of Experimental Psychology*, *70*(6), 1011–1022. https://doi.org/10.1080/17470218.2015.1101477

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., & Humphreys, G. W. (2017). 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

Vanes, L. D., White, T. P., Wigton, R. L., Joyce, D., Collier, T., & Shergill, S. S. (2016). Reduced susceptibility to the sound-induced flash fusion illusion in schizophrenia. *Psychiatry Research*, *245*, 58–65. https://doi.org/10.1016/j.psychres.2016.08.016

Verplanken, B., & Sui, J. (2019). Habit and identity: Behavioral, cognitive, affective, and motivational facets of an integrated self. *Frontiers in Psychology*, *10*, 1504. https://doi.org/10.3389/fpsyg.2019.01504

Wickham, H. (2016). ggplot2: Elegant graphics for data analysis. In *Journal of the Royal Statistical Society: Series A (Statistics in Society)* (Vol. 174, Issue 1, pp. 245–246). http://link.springer.com/10.1007/978-3-319-24277-4

Witt, J. K., Taylor, J. E. T., Sugovic, M., & Wixted, J. T. (2015). Signal detection measures cannot distinguish perceptual biases from response biases. *Perception*, *44*(3), 289–300. https://doi.org/10.1068/p7908

Witt, J. K., Taylor, J. E. T., Sugovic, M., & Wixted, J. T. (2016). Further clarifying signal detection theoretic interpretations of the Müller-Lyer and sound-induced flash illusions. *Journal of Vision*, *16*(11), 19–19. https://doi.org/10.1167/16.11.19