Title:

Introduction

The cognitive faculties responsible for understanding others and forming interpersonal relationships, known as social cognition, depend significantly on recognizing and decoding social cues, which include eye movements and facial expressions (Birmingham & Kingstone, 2009). Social attention has been investigated within this context using the Posner's paradigm (Posner, 1980), to explore the mechanisms of the spatial attentional orienting between social and non-social stimuli.

In this paradigm, a cue, like an arrow or gaze (i.e., non-social vs. social stimuli), signals the likely location of an upcoming target stimulus, leading to faster and more accurate responses when the target appears in the cued location (Friesen & Kingston, 1998). This phenomenon is referred to as the "cueing effect". Quantitatively, this effect has been consistent across both social and non-social cues (for a review, Chacón-Candia et al., 2022), suggesting that both types of stimuli share similar underlying spatial attentional mechanisms (Brignani et al., 2009; Corbetta & Shulman, 2002; Ristic & Kingstone, 2012). However, qualitative differences have been observed supporting the idea that gaze might have other specific dimension not shared with non-social cues (Marotta et al., 2012; Marotta et al., 2013; Chacón-Candia et al., 2020; Chacón-Candia et al., 2023).

One notable distinction was observed using the spatial interference paradigm (Lu & Proctor, 1995). When used as targets, arrows and eye-gaze elicit different conflict effects (Marotta et al., 2018). Arrows produced a standard congruency effect (SCE) with faster responses when direction and position were the same. Gaze cues, however, revealed a reversed congruency effect (RCE; Cañadas & Lupiáñez, 2012; Jones, 2015; Marotta et al., 2018), with faster responses during incongruent conditions (e.g., eyes looking to the right, presented on the left side of the screen).

This dissociation has been observed in multiple studies, offering new insights into potential social component implications and challenging certain hypotheses. Narganes-Pineda et al. (2022) demonstrated that explicit attention to the direction of eye-gaze and arrows is essential for specific effects to occur, challenging both the eye-contact hypothesis (Cañadas & Lupiáñez, 2012; Marotta et al., 2018) and the approach-avoidance motivational theory (Elliot, 2006). Meanwhile, Aranda-Martín and colleagues (2022) found that the reversed congruency effect is absent in children under 12, suggesting that it may not be related to joint attention (Edwards et al., 2020; Mundy, 2018). A "joint distraction" phenomenon has also been suggested as an explanation (Hemmerich et al., 2022), though recent experiments (Aranda-Martín et al., 2023) have not supported this theory. Nevertheless, the reversion effect seems to be influenced by the target face's emotional expressions (Jones, 2015; Torres-Martín et al., 2017; Marotta et al., 2022), and might be modulated by social familiarity (Ishikawa et al., 2022). There was also a negative correlation observed with social anxiety scores, a correlation that was not observed with arrows and word targets (Ishikawa et al., 2021). Additionally, the reversion, seems to be triggered by cropped-eyes, faces and even with inverted faces (Hemmerich et al., 2022; Román-Caballero et al., 2021a; Tanaka et al., 2022a), but not with another social orienting signal, such as pointing fingers (Bonventre & Marotta, 2023).

Moreover, both behavioural and neuroimaging evidence are consistent with the coexistence of a shared spatial interference mechanism between these stimuli, and an additional mechanisms that might lead to the reversion with social stimuli. Hemmerich et al., (2022) in addition to reproducing the SCE and the RCE for non-social and social stimuli respectively, also observed that arrows and words elicited the typical congruency sequence effect (CSE), while eye-gaze and faces resulted in a reversed CSE. More important, the RCE emerged with social and the SCE decreased with non-social stimuli after incongruent trials, independently of the preceding target type. In an ERP study, Marotta and colleagues (2019), observed that the stimuli caused opposite effects in late components N2 and P3, contrary to the early components P1 and N1 which showed no differences. Narganes-Pineda et al. (2023) in an fMRI study, observed similar activation in right parieto-temporo-occipital regions during conflict resolution for both stimulus types, whereas observed differential functional activation between the frontal eye field (FEF) and occipital regions.

On the other hand, studies involving perceptual manipulations have shown that the RCE is, to some degree, modulated by spatial coding processing related to task interference effects. In a prior study, Román-Caballero and colleagues (2021a), found a more pronounced RCE with faces under conditions where the SCE either reversed or was absent. For this, they used arrows embedded in a complex perceptual background resembling a face. It was, also noted that the RCE produced by whole faces was larger than the results from Marotta et al. (2018) who used cropped eyes. These findings imply that any factor reducing the spatial interference component, which is common to both social and non-social stimuli, would increase the reversal observed with gaze targets, as highlighted by the authors.

Similar results were observed in a follow-up study (Román-Caballero et al., 2021b) that included manipulations of synchrony between the background and the target presentation, giving synchronous and asynchronous experimental conditions. These manipulations revealed contrasting behaviours between the standard and the reversed effect: the SCE decreased in the synchronous condition, while the RCE decreased in the asynchronous condition. These modulations were evident even when arrows were embedded in a pixelated background with no facial resemblance. Furthermore, ongoing research from our laboratory (Román-Caballero et al., in progress, https://osf.io/a4nqh/), using peripheral cueing and alertness auditory signals, supports the notion that the SCE and the RCE are inversely related within specific experimental settings, and it is not modulated only at a perceptual level, but also with the implication of attentional processing.

Using comparable manipulations, Tanaka et al. 2022 reported similar patterns with tongue targets. Initially, they observed the SCE with arrows and the reversion with both gaze and tongue targets when using undistorted whole faces. Subsequently, the SCE and RCE were modulated using a perceptually complex background. They noticed a numerical reversion with arrows (though non-significant) and an increased reversion with tongue targets. From a different angle, Chen et al. (2022), by manipulating the stimulus perceptibility, found that when stimuli shared equivalent perceptual features, both social and non-social targets yielded the SCE (as seen in experiment 3).

These findings emphasize the significance of perceptual dimensions and the crucial role of cognitive processing time structures in conflict resolution with the social Stroop task. Román-Caballero et al (2021a; 2021b) argue that their results align with the temporal decay hypothesis (Hommel, 1993). This hypothesis postulates that the conflicting spatial dimension will have decayed by the time the relevant code is formed. In essence, the increased demands for target-

background segregation delay the processing of target-relevant information (such as the direction the target is pointing). This creates a temporal gap that leads to the reduction of SCEs with arrow targets. Consequently, by diminishing the spatial interference component with gaze targets, the RCE intensifies.

In a related manner, Tanaka et al. (2022), proposed a dual stage theory. Initially, the target-background segregation demand is facilitated (Román-Caballero et al. 2021a; 2021b). This delay subsequently activates a response inhibition process (Ridderinkhof, 2002), which induces a negative priming effect (Tipper, 1985). This effect results in slower responses in congruent conditions for gaze targets. However, for arrow targets, due to their perceptual simplicity, the responses occur before the onset of response inhibition, making responses faster in congruent conditions.

In summary, the spatial interference component, which seems to be common to both social and non-social stimuli, appears to need reduction for the reversion effect to manifest. This has been observed through manipulations of perceptual stimuli features (Román-Caballero et al., 2021a; 2021b) and on a trial-by-trial basis (Hemmerich et al., 2022). However, it is important to consider that achieving the reversion may not only require a reduction in the SCE but also an amplification of the RCE, that could be facilitated by a negative priming, but also by the phenomenon called inward-bias (Chen et al., 2018). This latter has been used to support the social component of the RCE hypothesis (Aranda-Martín et al., 2023; Román-Caballero et al., 2021a; 2021b).

The complex interplay between the SCE and RCE may be specific to certain experimental settings, while in others, they could exhibit unrelated behaviours, possibly due to the social component of eye-gaze. The aim of this analyses is to employ a distributional approach to explore the behaviour of the conflict effects within the distribution. This will help discern the similarities and differences between the SCE and RCE, by identifying specific portions of the distribution that might be uniquely influenced by trial congruency interactions. This insight might help us explore whether the reversion arises primarily from spatial processing aspects related to task characteristics and is later enhanced by the social significance of eye-gaze targets. Alternatively, it could be that the RCE is primarily driven by the social element and then further amplified by specific facets of the conflict task or perceptual processing. Such perspectives could be instrumental in comprehending the RCE and its interplay with the typical spatial interference.

METHOD

Datasets

We re-analyzed, adopting a distributional approach, three studies that employed the standard social Stroop task and two studies that used a perceptual background segregation task. The datasets were subjected to the same pre-processing procedures as in the original studies, which included the detection and removal of outliers. The table below provides details about the design and structure of these studies (for a statistical summary, see Table 1 in the supplemental material).

Table 1.

Study	N	Experiment Structure & Design	RT Filtering
Marotta et al. (2018)	35	- 2 halves, 1 by target type, 2 blocks each. 64 trials by block for a total of 256 - Within-subject design, 2 factors (Target type x Congruency)	200-1300 ms
Marotta et al. (2019)	27	- Same structure as 2018 - Variable duration of fixation	200-1300 ms
Hemmerich et al. (2022) - Exp 1	35	- Within-block target type design - Same structure and design as Marotta et al. (2018)	200-1300 ms
Hemmerich et al. (2022) - Exp 2	33	- Within-block target type design - Same structure and design as Marotta et al. (2019)	200-1300 ms
Hemmerich et al. (2022) - Exp 3	40	- Within-block target type design - 4 blocks, 512 trials in total - 2 factors, 4 stimulus types	200-1300 ms
Román-Caballero et al. (2021a) - Exp 1	35	- 2 blocks, 128 trials each - Same structure and design as Marotta et al. (2018)	±2.5 SD
Román-Caballero et al. (2021a) - Exp 2	35	- Same structura and design as experiment 1 - Change a feature in non-social stimuli (Figure 2B)	±2.5 SD
Román-Caballero et al. (2021b) - Exp 1	40	- Synchrony background-target manipulation - 4 blocks, 96 trials each (with synchrony variations) - Within-subject design, 3 factors (Synchrony x Target type x Congruency)	±2.5 SD
Román-Caballero et al. (2021b) - Exp 2	40	- Same design and structure as experiment 1 - Change social and non-social stimuli (see Figure 3)	±2.5 SD

Note. The number of participants, *N*, represents the sample size after outlier removal. The thicker line separates datasets from studies that employed the standard Stroop task and those that used the background segregation task

The stimuli used in the standard tasks are illustrated in Figure 1. Targets appear either to the left or right of a fixation cross. Participants were instructed to press the "z" for targets pointing to the left and the "m" for those pointing to the right, regardless of their position on the screen. In Hemmerich et al. (2022), a within-block design was employed, with target types presented randomly within the same block. In Experiment 3, in addition to arrows and eye-gaze, complete faces and words were introduced as stimuli. All experiments provided auditory feedback for incorrect responses and included 15 practice trials.

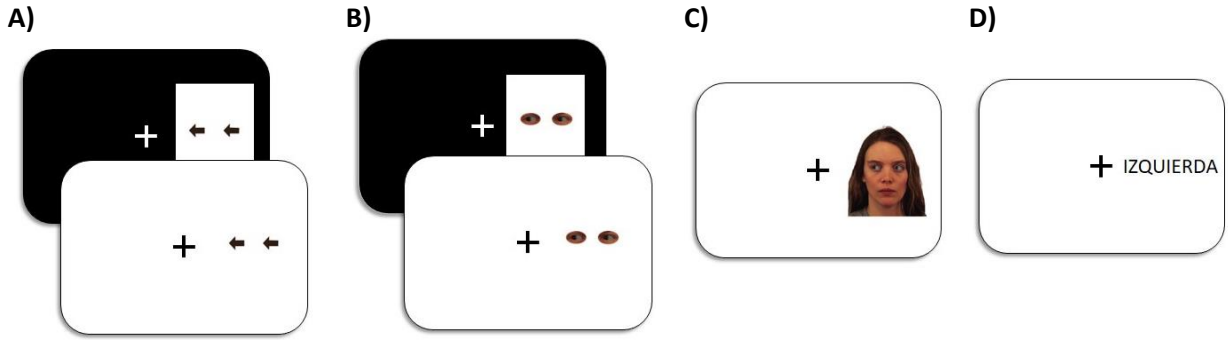


Figure 1. Example of the incongruent conditions in the experimental tasks. The stimuli from Marotta et al. (2018) were displayed against a black background embedded in a grey vertical rectangle (top (A) and (B)). In contrast, Marotta et al. (2019) presented their stimuli on a white background (bottom (A) and (B)). In Hemmerich et al. (2022), the black background version was employed in experiment 1, while the white background was used in experiment 2. Additionally, in experiment 3, face and word targets were introduced ((C) and (D)).

In Román-Caballero et al. (2021a), the perceptual background features for arrows were adjusted to heighten their complexity in processing while preserving a facial resemblance. Figure 2 illustrates this stimuli. Similar to the earlier task, auditory feedback was provided for incorrect responses, and participants were given 15 trials of practice.

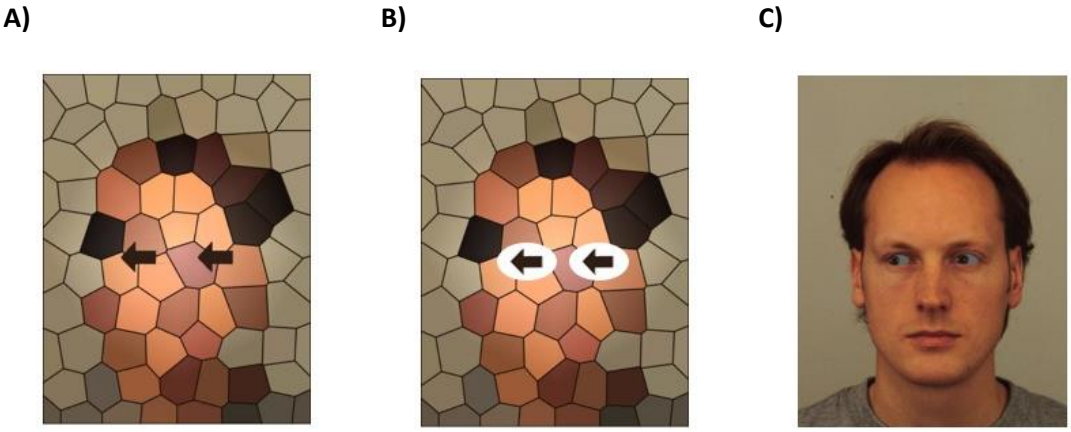


Figure 2. Example of stimuli used in Román-Caballero et al. (2021a). The arrows in (A) were used in experiment 1, whereas the arrows in (B) were used in experiment 2. The stimuli were presented in a black background either at the left or the right of a white fixation cross.

In Román-Caballero et al. (2021b), both experiments introduced a synchrony manipulation between the target and the background (synchronous vs. asynchronous). In the asynchronous block, the background was displayed without the targets followed by the target presentation. In the second experiment, for social stimuli targets, a face with closed eyes was presented during asynchronous condition before the target. An auditory feedback was used for incorrect responses.

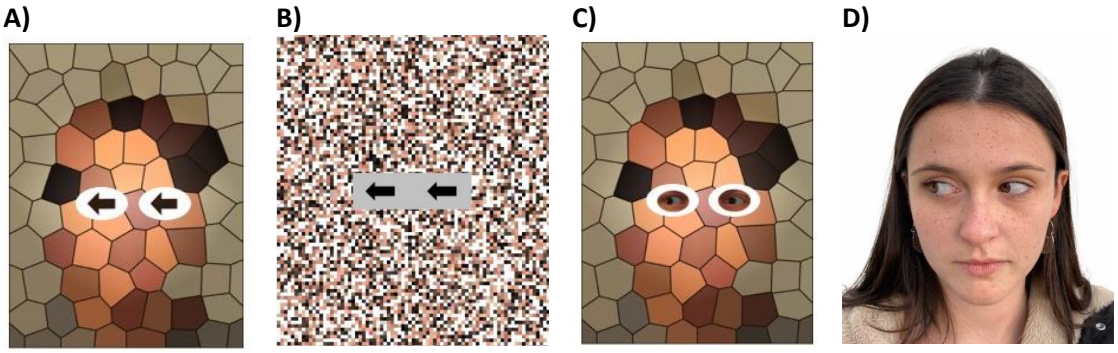


Figure 3. Stimuli used in Román-Caballero et al. (2021b), they were presented in a black background either at the left or at the right of the fixation cross. Stimuli (A) and (C) were used in experiment 1, while stimuli (B) and (C) were used in experiment 2.

Analyses

Analyses were conducted using R (R Core Team, 2023). The analysis scripts have been made publicly available via the Open Science Framework (OSF; LINK). Given the variability within and between participants, distributional analysis were performed using a mixed-model approach with the *lme4* (Bates et al., 2015) and *lmerTest* (Kuznetsova et al., 2017) packages. Additionally, we employed Bayesian analyses with the "BayesFactor" package (Morey & Rouder, 2022) to quantify the evidence for main effects and their interactions. In both approaches, participants were treated as random effects. Bonferroni post-hoc tests and Bayesian t-tests were used for specific comparisons. For the Bayesian analysis, we adhered to the default priors distributions and *rscale* parameters (Rouder et al., 2009 and Rouder et al., 2012). Bayes Factors (BF10) were interpreted based on the guidelines proposed by Raftery (1995; see also Wagenmakers, 2007). In one specific case, detailed below, we employed the *brms* package (Bürkner, 2017) to analyse the posterior distributions of the effect of interest.

Distributional Analysis

Distributional analysis examines the full spectrum of response times, rather than focusing solely on the average latencies (e.g., Heathcote et al., 1991; Mittelstädt & Miller, 2020; Tang et al., 2022). Essentially, this approach organizes reaction times from fastest to slowest and then segments them using quantile or Vincentile techniques. Such an approach provides insights into how an experimental effect unfolds throughout the response time distribution (Ratcliff, 1979; Balota & Yap, 2011). This technique has been extensively applied in conflict tasks (e.g., Mittelstädt & Miller, 2020; Miller & Schwarz, 2021) since it enables the temporal assessment of interactions between task-relevant and task-irrelevant dimensions. Three prevalent tools derived from these binned averages include the delta function, the conditional accuracy function (CAF), and the shift function.

The delta function illustrates the differences between two conditions (i.e., incongruent – congruent) against each bin to describe the interference effect's time course (Delta Plots; De Jong, 1994). We employed Vicentization (Ratclif, 1979; Vincent, 1912) on correct responses to divide the RT distribution into five equally-sized bins per participant and target type. Using an orthogonal-polynomial-contrast analysis, which helps assess the shape of the delta plots (Burle et al., 2014; Pratte, 2021), we quantified each target's general trend. A mixed-model approach with participants as random effects was adopted. This analyses determined whether the interference effect followed a linear or quadratic general trend. Additionally, we analyzed the slopes segments (similarly to Debey et al., 2015 & Ridderinkhof et al., 2005) to evaluate significant distributional differences concerning each portion. One-sample t-tests against zero were used for this purpose.

At a distributional level, we also investigated the behaviour of the congruency sequence effects (CSEs). The CSE can be quantified as: $CSE = RTs_{[(cI - cC) - (iI - iC)]}$, where *cC* and *cI* are congruent and incongruent trials, respectively, preceded by congruent trials; and *iC* and *iI* are congruent and incongruent trials following incongruent trials (Nieuwenhuis, et al., 2006; Tang et al., 2022). For this analysis, we considered only correct responses from both the preceding and current trials, excluding the first trial of each block. Hemmerich et al. (2022) employed a within-block design study, in which they also took into account the sequence of target type trials, resulting in conditions for both the same and different previous target type trials. Upon examining this design, we opted not to use this categorization due to a significant reduction in the number of

trials per bin for each participant. For the same reason we did not include the third experiment in this analysis. To test the effect size of the two possible sequential conflict effects (*cIcC* and *iIiC*) we used one sample t-tests against zero.

The CAFs represent response accuracy as a function of response speed, offering a dynamic perspective of accuracy across varying reaction times (e.g., Ambrosi et al., 2019; Ülrich et al., 2015). Like delta functions, we analysed our datasets using Vincentization but considered both correct and incorrect responses per participant. This resulted in a three-factor within-subject design: target type (with four targets in experiment 3 of Hemmerich et al. 2022), congruency, and bin (2/4 x 2 x 5). Post-hoc Bonferroni tests and paired Bayesian t-tests were conducted to compare congruency levels in each bin.

To better understand the influence of accuracy rate on conflict effects, we assessed congruency effects (i.e. incongruent – congruent trials) using both the linear integrated speed-accuracy score (LISAS; Vandierendonck, 2017) and mean RTs for each participant's condition. The LISAS represents an error-adjusted RT measure that incorporates the standard deviations of correct responses and proportion of errors, and thus provided with a weighted score for reaction times (Vanderendonck, 2021). For instance, if a participant obtain an average correct RT of 600 with a standard deviation of 100, and an error rate of .10 with a standard deviation of .06 for a specific condition, the computed LISAS would be 766.67.

Significantly, numerous studies (e.g., Marotta et al., 2018, 2019; Aranda-Martín et al., 2022; Narganes-Pineda et al., 2022; Bonventre & Marotta, 2023) have observed that arrows yield faster response times compared to eye-gaze and face targets. Thus, it was considered important to analyse this relationship at a distributional level. The shift function, which employs the Harrell-Davis quantile estimator (Harrel & Davis, 1982) and typically segments the distribution into deciles (Wilcox, 1995), estimates the magnitude and way by which one distribution must be adjusted to match another (Doksum, 1974; Rousselet et al., 2017). We used only correct responses and calculated the decile differences between eye-gaze or face and arrows targets. To test if the differences were significant from zero, high density intervals (HDI) were computed using Bayesian bootstrapping (Rubin, 1981) via the R package *bayesboot* (Bååth, 2022).

RESULTS

Congruency effect Delta Plots

The orthogonal-polynomial contrasts revealed significant linear coefficients for both arrows and eye-gaze in Marotta et al. (2018) as well as in experiment 3 of Hemmerich et al. (2022). Likewise, for both target types, the linear term was significant in experiments 1 and 2 of the synchronous block in Román-Caballero et al. (2021b), and in experiment 2 of the asynchronous block. These results imply that the conflict effect decreases for arrows while it increases for social targets across the distribution. Comparable outcomes were observed with social targets in both experiments of Román-Caballero et al. (2021a). These findings remained significant when using 4 bins, with the exception of experiment 2 in the asynchronous block from Román-Caballero et al. (2021a).

Table 1. Orthogonal Polynomial Contrasts of Linear and Quadratic Coefficients.

Study		Linear	Quadratic

	Target Type	*F*	*Df*	*p*	*F*	*Df*	*p*
Marotta et al. (2018)	Arrows	23.93	1, 138	**<.0001**	1.86	1, 138	.175
	Eye-Gaze	23.98	1, 138	**<.0001**	7.61	1, 138	**.007**
Marotta et al. (2019)	Arrows	1.80	1, 106	.182	0.00	1, 106	.975
	Eye-Gaze	2.78	1, 106	.099	.75	1, 106	.389
Hemmerich et al. (2022, exp. 1) **+**	Arrows	.20	1, 138	.657	5.64	1, 138	**.018**
	Eye-Gaze	1.39	1, 138	.241	12.70	1, 138	**<.001**
Hemmerich et al. (2022, exp. 2) **+**	Arrows	.83	1, 130	.365	8.53	1, 130	**.004**
	Eye-Gaze	1.80	1, 130	.183	15.71	1, 130	**<.001**
Hemmerich et al. (2022, exp. 3) *****	Arrows	7.99	1, 158	**.005**	.04	1, 158	.837
	Eye-Gaze	4.21	1, 158	**.041**	6.09	1, 158	**.015**
	Word	.10	1, 158	.755	.84	1, 158	.362
	Face	1.85	1, 158	.175	.06	1, 158	.80
Román-Caballero et al. (2021a, exp. 1) **+**	Arrows	.73	1, 138	.395	11.27	1, 138	**.001**
	Face	24.19	1, 138	**<.0001**	3.79	1, 138	.054
Román-Caballero et al. (2021a, exp. 2)	Arrows	.43	1, 138	.513	7.28	1, 138	**.008**
	Face	9.58	1, 138	**.002**	2.65	1, 138	.106
Román-Caballero et al. (2021b, exp. 1, sync.) **+**	Arrows	16.91	1, 158	**<.001**	19.00	1, 158	**<.0001**
	Eye-Gaze	11.64	1, 158	**<.001**	35.59	1, 158	**<.0001**
Román-Caballero et al. (2021b, exp. 1, async.)	Arrows	1.02	1, 158	.314	.00	1, 158	.991
	Eye-Gaze	36.53	1, 158	**<.0001**	2.52	1, 158	.114
Román-Caballero et al. (2021b, exp. 2, sync.) *	Arrows	13.12	1, 158	**<.001**	4.00	1, 158	**.047**
	Face	28.85	1, 158	**<.0001**	4.41	1, 158	**.037**
Román-Caballero et al. (2021b, exp. 2, async.) *****	Arrows	4.30	1, 158	**.04**	1.34	1, 158	.25
	Face	59.32	1, 158	**<.0001**	3.24	1, 158	.074

Note. (*****) relevant changes when using 4 bins: in Hemmerich et al. (2022, experiment 3), the quadratic term with eye-gaze targets was no significant *(p = .066*). In Román-Caballero et al. (2021b), in the synchronous condition of the experiment 2, the two quadratic terms were no longer significant (*p = .149* & *p = .123*, respectively), whereas in the synchronous block, the linear term for arrows was *p =.062*. (**+**) relevant changes of the quadratic term when using 3 bins: the quadratic term remained significant for eye-gaze in Hemmerich et al. (2022) in experiments 1 and 2, and for arrows in experiment 1 in Román-Caballero et al. (2021a). In Román-Caballero et al. (2021b), in the synchronous condition of experiment 1 remained significant for both target types.

In Marotta et al. (2018), the quadratic term was significant for the eye-gaze target using both 4 and 5 bins. In experiments 1 and 2 of Hemmerich et al. (2022), this significance persisted for both targets using 4 and 5 bins and was also observed for the eye-gaze targets when employing 3 bins. Similarly, in experiment 1 of Román-Caballero et al. (2021b), the quadratic term remained significant in the synchronous condition with 3, 4, and 5 bins for both target types. In contrast, in experiments 1 and 2 of Román-Caballero et al. (2021a), the quadratic term remained significant for arrow targets with 4 and 5 bins and continued to be significant in experiment 1 using 3 bins.

The presence of a significant quadratic term suggests that for arrow targets, conflict effects decrease with faster responses and increase with slower ones. For eye-gaze or face targets, the pattern is reversed, implying an increment of the conflict effect with faster responses. Specifically, the within-block condition seems to favour a U-shaped trend for eye-gaze. When targets are presented in separate blocks, this trend appears to be exclusive to arrow targets embedded in backgrounds resembling faces, as seen in experiments 1 and 2 of Román-Caballero et al. (2021a), whereas when complete faces were utilized the quadratic term was not significant. The only significant quadratic term for the eye-gaze with background segregation

tasks target was observed in the synchronous condition of experiment 1 in the 2021b study. In Figure 4 it is possible to observe the patterns in Román-Caballero et al. (2021b) study.

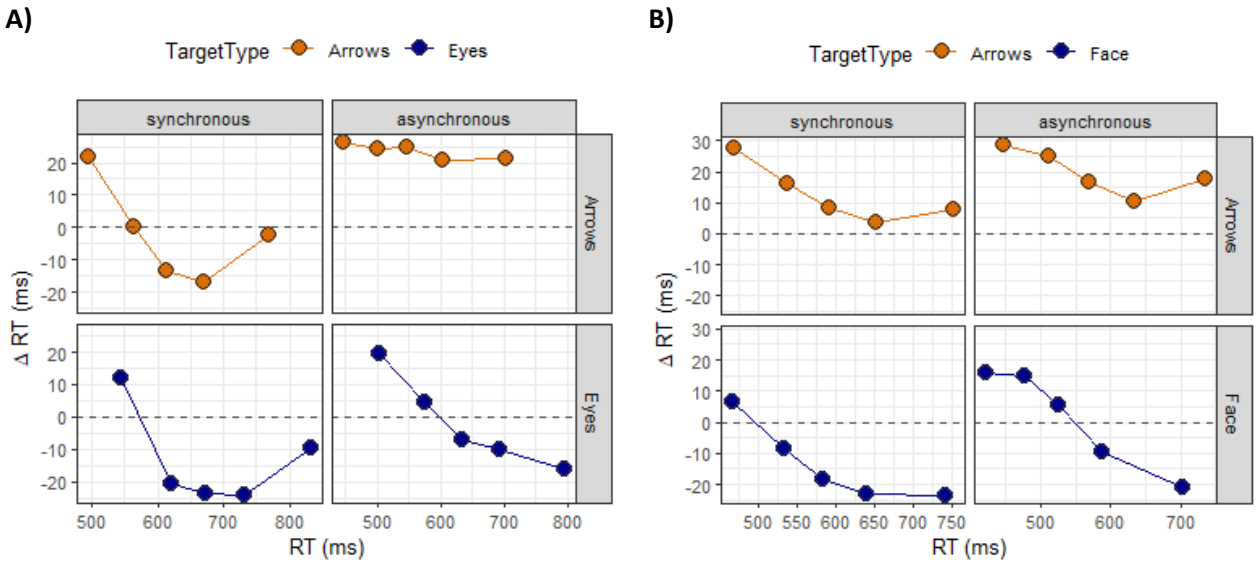


Figure 4. Delta plots of experiment 1 (A) and experiment 2 (B) of Román-Caballero et al. (2021b) study. Delta plots display the delta values (i.e., the subtraction of congruent trials from incongruent trials, they represent the effect size) by bin, plotted against the mean of that bin.

Using a Bayesian approach, datasets with 4 and 5 bins were compared. Figure 1 of the supplemental material shows the distributions of posterior samples for the quadratic terms of the orthogonal contrast model. The critical studies showed that the distributions are situated mostly to the right of zero, indicating the predominance of this pattern for eye-gaze in experiments 1 and 2 of Hemmerich et al. (2022) and in experiment 1 of Román-Caballero et al. (2021b). This pattern is also evident for arrow targets in experiments 1 and 2 of the Román-Caballero et al. (2021a) study.

Regarding the segment slopes of the delta plots, they suggest that the critical changes in the distributions with respect to the conflict effect are most evident within the range of the fastest responses, as depicted in Table 2 (for complete results, refer to Table 2 in the supplemental material). In general, there were significant negative-going slopes for both targets, especially in the B1-2 segment. This indicates a rapid decrease with arrows and an increase with social targets concerning the conflict effect. Such a trend was not observed in experiment 3 of Hemmerich et al. (2022) with arrows, words, and face targets, implying that the effects remained constant throughout the distribution. Notably, in Román-Caballero et al. (2021b) experiment 2 of the asynchronous condition, the negative nature of the slopes began to become significant from the B2-3 segment, and continued to be significant up to the B4-5 segment for face targets, reaching completely the reversion. This suggests that the reversal of the effect is more pronounced for median and slower response times in this condition.

Table 2. Delta plot slopes analysis of B1-2 and B2-3 segments.

Study	Target Type	B1-2		B2-3	
		p	B_{10}	*p*	B_{10}
Marotta et al. (2018)	Arrows	**.003**	**13.33**	**<.0001**	**>1000**
	Eye-Gaze	**<.001**	**>1000**	**.009**	**4.59**

Marotta et al. (2019)	Arrows	.563	.238	**.016**	**3.20**
	Eye-Gaze	.113	.663	**.007**	**6.20**
Hemmerich et al. (2022, exp. 1)	Arrows	**<.0001**	**>1000**	.465	.234
	Eye-Gaze	**<.001**	**256.4**	.106	.626
Hemmerich et al. (2022, exp. 2)	Arrows	**.003**	**13.06**	.279	.324
	Eye-Gaze	**.004**	**9.13**	.528	.225
Hemmerich et al. (2022, exp. 3)	Arrows	.088	.685	.197	.377
	Eye-Gaze	**<.0001**	**>1000**	.084	.710
	Word	.119	.546	.113	.566
	Face	.267	.307	.205	.367
Román-Caballero et al. (2021a, exp. 1)	Arrows	**.003**	**13.34**	.193	.406
	Face	**<.001**	**235.3**	**.015**	**3.037**
Román-Caballero et al. (2021a, exp. 2)	Arrows	**.002**	**19.09**	.771	.189
	Face	**.002**	**17.06**	**.003**	**11.28**
Román-Caballero et al. (2021b, exp. 1, sync.)	Arrows	**<.0001**	**>1000**	**<.0001**	**>1000**
	Eye-Gaze	**<.0001**	**>1000**	.272	.303
Román-Caballero et al. (2021b, exp. 1, async.)	Arrows	.746	.179	.922	.171
	Eye-Gaze	**<.001**	**69.56**	**.002**	**14.19**
Román-Caballero et al. (2021b, exp. 2, sync.)	Arrows	**.007**	**5.89**	**.017**	**2.59**
	Face	**<.001**	**69.68**	**<.001**	**35.15**
Román-Caballero et al. (2021b, exp. 2, async.)	Arrows	.323	.272	**.004**	**8.90**
	Face	.608	.193	**<.001**	**88.48**

Note. The slopes were calculated by participant using the common linear formula, slope = (y2 – y1)/(x2 – x1) where the values of y-axis are the delta values (incongruent – congruent trials), and the x-axis values are the mean RT of each bin. B1-2 indicates the slope of the segment from bin 1 to bin 2.

About the last segment B4-5, it displayed significant positive-going slopes for social targets in Hemmerich et al. (2022, experiment 2) and Román-Caballero et al. (2021b, experiment 1) in the synchronous condition, with positive evidence in favour (BF10 = 4.11 & BF10 = 4.83 respectively). Although it was not statistically significant in experiment 1 of Hemmerich et al. (2022), a numerically positive-going slope in B4-5 was still observed (Figure 5).

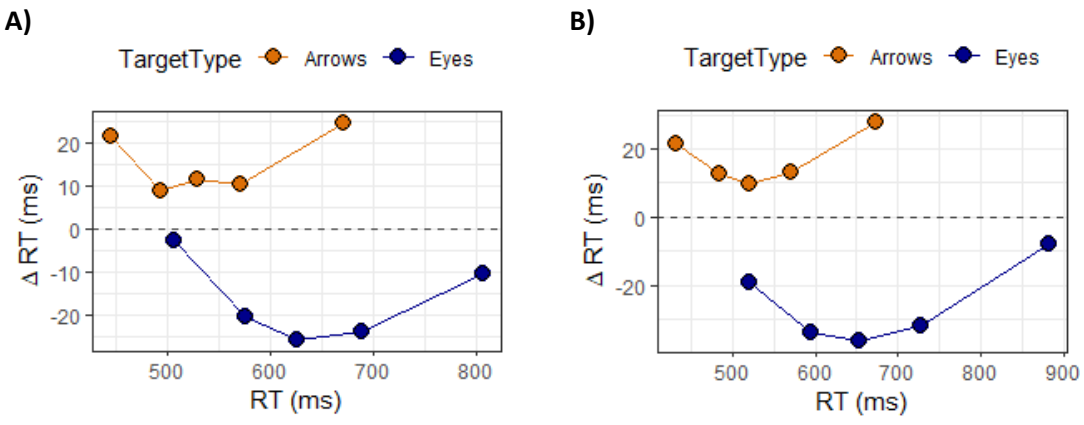


Figure 5. Delta plots of experiment 1 (A) and experiment 2 (B) of Hemmerich et al. (2022).

Congruency sequence effect Delta Plots

The sequential analysis results indicated that the reversions with social targets were facilitated in the *iIiC* condition (i.e., the congruency effect when previous trials were incongruent). This pattern was observed throughout the distribution in every standard task study. Moreover, the one sample t-tests revealed significant differences from zero, showing strong evidences in favour in most of the cases (BF10 > 100; see Table 3 in the supplemental material), except in bin 5 of experiment 1 in Hemmerich et al. (2022). For the *cIcC* condition (i.e., congruency effect when previous trials were congruent) with social targets, a significant SCE was observed that rapidly decayed, lasting from bin 1 to 2 (Marotta et al., 2018), or just for Bin 1 (experiment 1 in Hemmerich et al., 2022), as depicted in Figure 6. This was followed by a non-significant reversion. In contrast, in Marotta et al. (2019) showed no significant SCE initially, followed by a non-significant reversion.

For arrow targets, in the *iIiC* condition, a reversion was observed in certain segments of the distribution, with only in a few bins being significantly different from zero (see Table 3 in the supplemental material). In *cIcC* condition, however, the SCE was significant across the entire distribution.

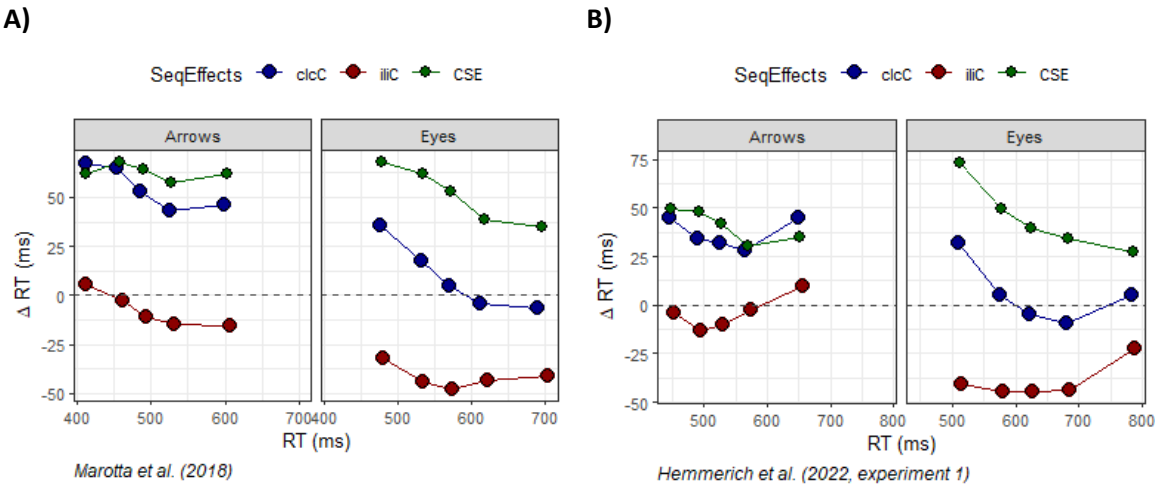


Figure 6. Delta plots of the congruency sequence effects of Marotta et al. (2018) (A), and experiment 1 of Hemmerich et al. (2022) (B). *cIcC* represents the effect size for trials that follow congruent ones, while *iIiC* denotes the effect size for trials following incongruent ones. The congruency sequence effect (*CSE*) was calculated as *cIcC - iIiC*.

With respect to the background segregation tasks, in Román-Caballero et al. (2021a) it was observed a reversion with both target types in the *iIiC* condition across the distribution. This pattern also emerged in the synchronous block of experiment 1 in Román-Caballero et al. (2021b), where backgrounds were similar to those used in their 2021a study. Notably, this trend was not evident for arrow targets in the other experiments and conditions, where the effect size oscillated around zero. Moreover, in the asynchronous condition of the 2021b study, a reversion was observed in both experiments with social targets. However, only in experiment 1, which used cropped eyes (Figure 3C), it was statistically different from zero (see Table 3 in the supplemental material).

In the *cIcC* condition, a reversion was observed in both experiments of Román-Caballero et al. (2021a) with social targets. Specifically, in experiment 1, a strong evidence emerged in bin 4

(BF10 = 104.9), while in experiment 2, a positive evidence was seen in bin 3 (BF10 = 5.40). Whereas, in the 2021b study, the SCE was significant, especially with faster responses, followed by a rapid decay, most notably in the synchronous block (Figure 7). For arrow targets, in experiment 1 of the 2021a study, the pattern was similar to standard Stroop tasks studies, with social targets indicating a rapid decay of the SCE (only significant in the first bin). However, when adding the white ovals within the target presentation (Figure 2B), the SCE lasted longer, showing no decay throughout the distribution. These patterns were observed in Román-Caballero et al. (2021b); the synchronous condition elicited a rapid decay of the SCE (bin 1 and 2), whereas in the asynchronous condition, the decay occurred around bins 4 and 5.

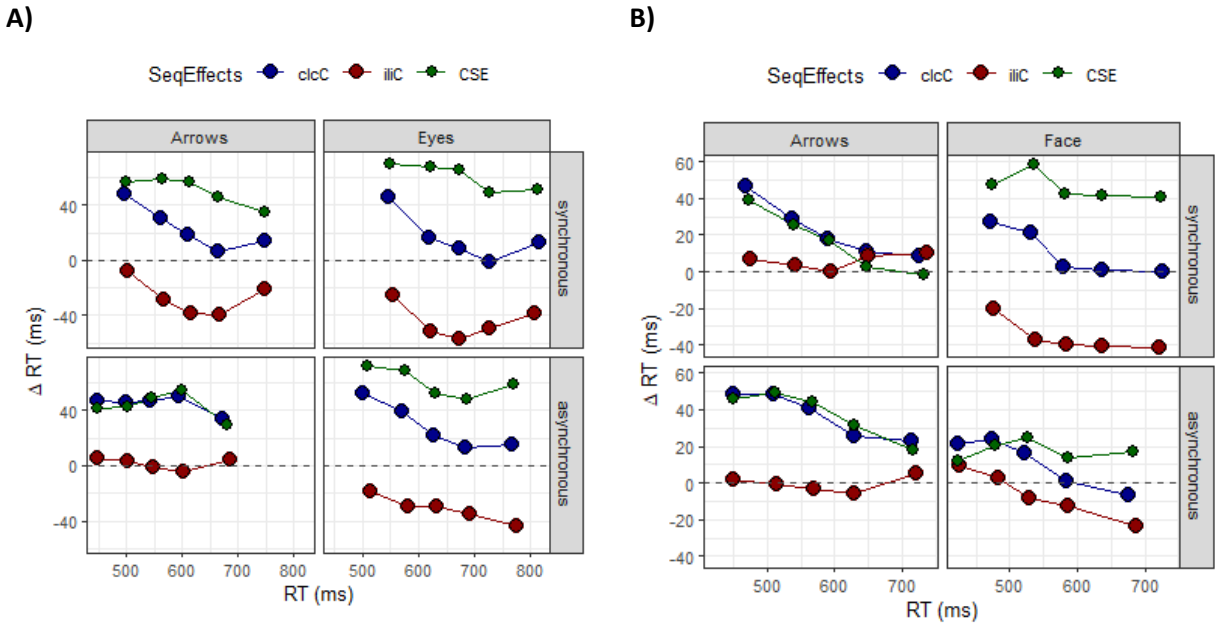


Figure 7. Delta plots of the congruency sequence effects of Román-Caballero et al. (2021b, experiment 1) (A), and Román-Caballero et al. (2021b, experiment 2) (B). *cIcC* represents the effect size for trials that follow congruent ones, while *iIiC* denotes the effect size for trials following incongruent ones. The congruency sequence effect (*CSE*) was calculated as *cIcC* - *iIiC*.

Conditional Accuracy Functions (CAFs)

Results from CAFs highlighted certain modulations in accuracy rate influenced by congruency. Significant three-way interactions were observed in the standard Stroop task studies, with the exception of Hemmerich et al. (2022), experiment 1. In Marotta et al. (2018) the evidence supporting the three-way interaction was weak (BF10 = 2.66), whereas the other studies presented positive evidence (BF10s > 3). For the background segregation studies, neither the three-way interactions in Román-Caballero et al. (2021a), nor the four-way interactions in Román-Caballero et al. (2021b) were significant (see Table 4 in the supplemental material).

The post-hoc Bonferroni tests and the paired Bayesian t-tests (Table 3. The complete results are reported in Table 5 in the supplemental material), either indicated weak evidence or did not support the reversions observed for social targets across faster responses in the CAFs plots (see Figure 4 in the supplemental material), with the standard Stroop tasks. Notably, in the asynchronous condition of the 2021b study, positive evidence emerged in experiment 1 (BF10 =

3.23), and strong evidence was evident in experiment 2 for bin 1 (BF10 = 38.68), suggesting however, a standard effect, i.e., higher accuracy rates for congruent compared to incongruent trials for bin 1s (Figure 8).

Table 3. p-values of the post-hoc Bonferroni tests and Bayes factors of the paired Bayesian t-tests.

Study	Target Type	Bins			
		1		2	
		p	BF_{10}	*p*	BF_{10}
Marotta et al. (2018)	Arrows	**<.001**	**67.14**	.24	.35
	Eye-Gaze	.82	.19	.15	.50
Marotta et al. (2019)	Arrows	**.002**	**16.70**	.82	.21
	Eye-Gaze	.15	.54	.34	.31
Hemmerich et al. (2022, exp. 1)	Arrows	.52	.22	.95	.18
	Eye-Gaze	.14	.51	.06	.92
Hemmerich et al. (2022, exp. 2) *	Arrows	.052	**1.12**	.09	.71
	Eye-Gaze	**.04**	**1.30**	.88	.19
Hemmerich et al. (2022, exp. 3)	Arrows	**.02**	**2.94**	.97	.17
	Eye-Gaze	**.03**	**1.80**	**.05**	**1.01**
	Word	**.002**	**21.01**	.08	.72
	Face	.93	.20	.71	.18
Román-Caballero et al. (2021a, exp. 1)	Arrows	.67	.20	.60	.21
	Face	.07	.83	**.03**	**1.56**
Román-Caballero et al. (2021a, exp. 2) *	Arrows	.38	.26	.97	.18
	Face	.08	.75	**.04**	**1.48**
Román-Caballero et al. (2021b, exp. 1, sync.)	Arrows	.08	.72	.59	.20
	Eye-Gaze	.28	.30	.38	.25
Román-Caballero et al. (2021b, exp. 1, async.)	Arrows	**<.0001**	**>1000**	.04	**1.21**
	Eye-Gaze	**.01**	**3.23**	.64	.19
Román-Caballero et al. (2021b, exp. 2, sync.)	Arrows	**.004**	**8.58**	.82	.17
	Face	.69	.18	.08	.70
Román-Caballero et al. (2021b, exp. 2, async.)	Arrows	**<.0001**	**>900**	.36	.25
	Face	**<.001**	**38.68**	.67	.19

Note. Significant p-values and Bayes factors in favour of the alternative hypothesis are in bold. The asterisks (*) indicates whether differences were observed using 4 Bins. In Hemmerich et al. (2022, exp. 2) with arrow target it is obtained a significant value (*p = .025*) in bin 1. In Román-Caballero et al. (2021a, exp. 2) was observed no significant effects in the first two Bins for Face targets (*p = .076 & p = .151*). Significant results are in bold.

Conversely, for non-social targets, significant findings revealed both positive and strong evidence of decreased accuracy rates for incongruent trials in bin 1, in Marotta et al., (2018; 2019), and in experiment 3 of Hemmerich et al. (2022) study. A similar pattern was observed in Román-Caballero et al. (2021b) in the asynchronous condition of experiment 1, and in both synchrony conditions in the experiment 2.

A) **B)**

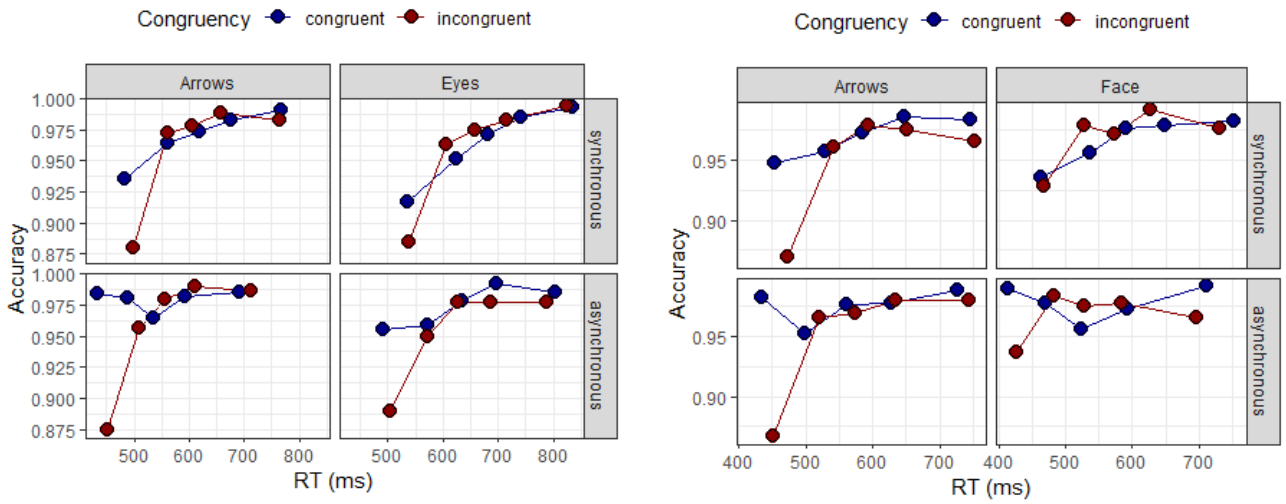


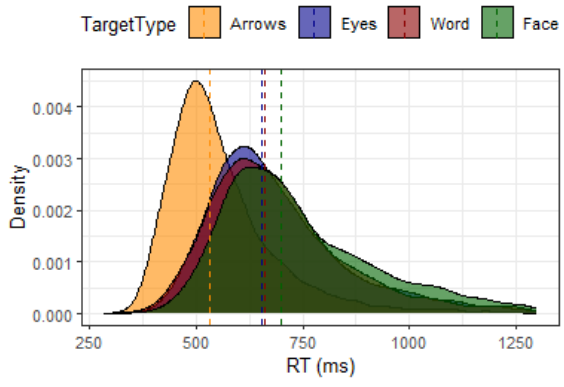
Figure 8. CAFs of the Experiment 1 (A), and 2 (A) of Román-Caballero et al. (2021b).

When we employed the LISAS for analysis, it highlighted specific modulations with arrow targets in certain studies (see Table 7 in the supplemental material). By also taking incorrect responses into account, the conflict effects increased. Both Marotta et al (2019), and experiment 2 of Hemmerich et al. (2022) demonstrated strong and positive evidence supporting a distinction between LISAS and mean RTs (BF10 = 73.96 & BF10 = 7.61). This was also observed in both experiments of the asynchronous condition in Román-Caballero et al. (2021b).

Shift Function

As expected the distributions of social targets exhibited a significant rightward shift compared to the arrows distribution. This is due to the faster reaction times associated with arrow targets in standard Stroop task studies (see, for instance, Figure 9). whit arrow targets in standard stroop task studies (e.g., Figure 9). This shifting was evident across the entire distribution, but it becomes more pronounced with slower responses.

A)



B)

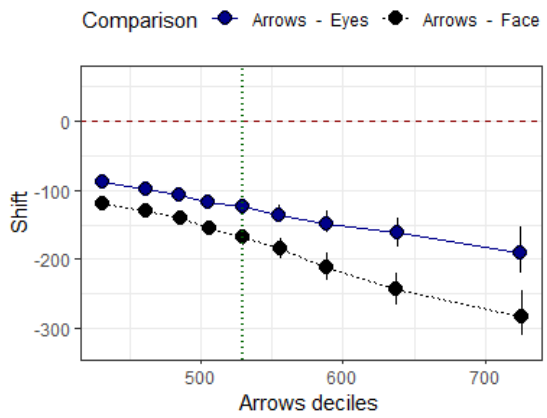


Figure 9. Kernel density estimation (A) and shift function (B) of Hemmerich et al. (2022, experiment 3). The shift values for each social target were determined relative to the arrow targets and are plotted against the mean RT for each decile of the arrows. The green dashed line cross the 0.5 decil (median) for the arrows. Bars represent the high density intervals (HDI) computed using bayesian bootsrapping.

Conversely, in studies employing the background segregation tasks, modulations in reaction times emerged in specific segments of the distribution. In Román-Caballero et al. (2021a) experiment 1, the most pronounced effect was observed at slower response times. In contrast, in experiment 2, the two distributions were closely aligned. (Figure 5 in the supplemental material). Notably, this pattern did not reappear in experiment 1 of the 2021b study, in either the synchronous or asynchronous conditions. With the same perceptual background for both stimuli, arrow distributions shifted to the left compared to eye-gaze. Furthermore, when comparing arrows embedded in a pixelated background with face targets, the distributions were similar in the synchronous condition. However, in the asynchronous condition, the reaction times for arrows became significantly slower than those for face targets (Figure 10).

A) **B)**

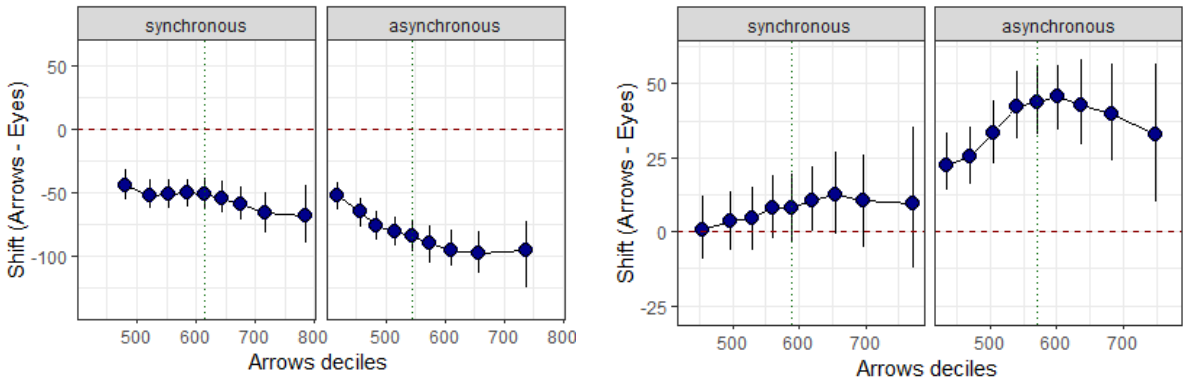


Figure 10. Shift functions of Román-Caballero et al. (2021b) study: A) experiment 1 and B) experiment 2. The shift values are plotted against the mean RT of arrows for each decile. The green dashed line marks the 0.5 decile (median) of the arrows. Bars represent the high density intervals (HDI) computed through Bayesian bootsrapping.

DISCUSSION

- Gaze cues facilitate the processing in directing attention to socially relevant information (Yuan et al., 2022). And affective contexts are also shown to modulate the gaze cueing effect (Lassalle et al., 2015; Zhang et al., 2021). In addition, also facilitates the processing of gazed targets in memory (Gregory & Jackson, 2017)
- Although arrows were slower in Román-Caballero et al. (2021b, experiment 2), suggesting that they need more time in extracting the relevant information from the background. Thus, making it possible for the response inhibition to appear, there were no reversion for arrows neither in *iIiC* condition nor with congruency effect delta plots.
- Si es top down lo de tipper significa que en ese momento puede actuar lo social de la mirada→ si apunta a lo social relevante entonces tiene que existir una intención mas alla de lo perceptivo ?¿

- Ulrich suppression has been taken to be more efficient for slower responses. More specifically, this is often indicated by negative-going delta functions in the visual Simon task and by initially positive-going and subsequently flat delta functions in theEriksen flanker task -> in gaze the suppression is for the SCE
- ERP LRP se active la parte motriz solo con pensar en la respuesta que debería presionar
- N170 es mayor con ojos? -> Eso quiere decir que el procesamiento de rostros es inicial es decir cuando se inicia con la segregación del background el sistema ya sabe que se trata de un rostro,

REFERENCES