

Seeing Social: A Neural Signature for Conscious Perception of Social Interactions

 Rekha S. Varrier and  Emily S. Finn

Department of Psychological and Brain Sciences, Dartmouth College, Hanover, New Hampshire 03755

Social information is some of the most ambiguous content we encounter in our daily lives, yet in experimental contexts, percepts of social interactions—that is, whether an interaction is present and if so, the nature of that interaction—are often dichotomized as correct or incorrect based on experimenter-assigned labels. Here, we investigated the behavioral and neural correlates of subjective (or conscious) social perception using data from the Human Connectome Project in which participants ($n = 1049$; 486 men, 562 women) viewed animations of geometric shapes during fMRI and indicated whether they perceived a social interaction or random motion. Critically, rather than experimenter-assigned labels, we used observers' own reports of "Social" or "Non-social" to classify percepts and characterize brain activity, including leveraging a particularly ambiguous animation perceived as "Social" by some but "Non-social" by others to control for visual input. Behaviorally, observers were biased toward perceiving information as social (vs non-social); and neurally, observer reports (compared with experimenter labels) explained more variance in activity across much of the brain. Using "Unsure" reports, we identified several regions that responded parametrically to perceived socialness. Neural responses to social versus non-social content diverged early in time and in the cortical hierarchy. Finally, individuals with higher internalizing trait scores showed both a higher response bias toward "Social" and an inverse relationship with activity in default mode and visual association areas while scanning for social information. Findings underscore the subjective nature of social perception and the importance of using observer reports to study percepts of social interactions.

Key words: social perception; individual differences; visual; animations; fMRI; internalizing traits

Significance Statement

Simple animations involving two or more geometric shapes have been used as a gold standard to understand social cognition and impairments therein. Yet, experimenter-assigned labels of what is social versus non-social are frequently used as a ground truth, despite the fact that percepts of such ambiguous social stimuli are highly subjective. Here, we used behavioral and fMRI data from a large sample of neurotypical individuals to show that participants' responses reveal subtle behavioral biases, help us study neural responses to social content more precisely, and covary with internalizing trait scores. Our findings underscore the subjective nature of social perception and the importance of considering observer reports in studying behavioral and neural dynamics of social perception.

Introduction

A remarkable feature of human perception is how quickly and automatically we identify social information in the environment. Consider pareidolia (seeing illusory faces in everyday objects; Palmer and Clifford, 2020) or the cocktail party effect (perceiving

self-relevant cues in otherwise unattended information streams; Wood and Cowan, 1995).

In the brain, the superior temporal sulcus (STS) has been classically associated with social cognition. Posterior STS regions are involved in perceiving animacy (Lee et al., 2014) and determining the nature of interactions (Isik et al., 2017), whereas anterior regions are involved in mentalizing, language, and gaze detection (Deen et al., 2015). Social signal detection, however, may begin even earlier in the lateral occipital and inferotemporal regions where recent work has proposed Gestalt-like perceptual mechanisms for social content (i.e., grouping social entities like facing dyads; Walbrin and Koldewyn, 2019; Abassi and Papeo, 2020; Papeo, 2020; Landsiedel et al., 2022). The recently proposed third visual stream (Pitcher and Ungerleider, 2021) posits a specialized pathway for social information that connects primary visual cortex (V1) to the motion-processing region V5/MT and culminates in the STS.

Received May 4, 2022; revised Aug. 15, 2022; accepted Oct. 14, 2022.

Author contributions: R.S.V. and E.S.F. designed research; R.S.V. performed research; R.S.V. and E.S.F. analyzed data; R.S.V. and E.S.F. wrote the paper.

This work was supported by a National Association for Research on Schizophrenia and Depression (Brain and Behavior Research Foundation) Young Investigator Award (Grant 28392 to E.S.F.), Dartmouth College Neukom CompX Faculty Grant (E.S.F.), and National Institute of Mental Health Grant R01MH129648 (E.S.F.).

The authors declare no competing financial interests.

Correspondence should be addressed to Rekha S. Varrier at Rekha.S.Varrier@dartmouth.edu or Emily S. Finn at emily.s.finn@dartmouth.edu.

<https://doi.org/10.1523/JNEUROSCI.0859-22.2022>

Copyright © 2022 the authors

Our tendency to spontaneously perceive social interactions in simple animations of geometric shapes emphasizes the relevance of motion to social perception (Heider and Simmel, 1944; Scholl and Tremoulet, 2000). This phenomenon transcends age and culture (Barrett et al., 2005; Mohammadzadeh et al., 2012) but perhaps not species (Schafroth et al., 2021). Although robust, percepts of these animations do vary across individuals. People with autism are less likely to report social interactions (Abell et al., 2000) and show commensurately lower brain activity in social processing regions (Castelli et al., 2002; Kana et al., 2015). Even neurotypical individuals differ in their socioperceptual tendencies (Rasmussen and Jiang, 2019; Li et al., 2020) in ways that covary with traits like loneliness, anxiety, and autism-like phenotypes (Kanai et al., 2012; Powers et al., 2014; Lisoy et al., 2022).

Past work has largely used stimuli handcrafted to be perceived as social or non-social and relied on these experimenter-assigned labels to contrast behavior and/or brain activity. Consequently, effects could reflect not only responses to social information but also differences across animations in basic physical properties (e.g., speed), which are rarely systematically controlled. Further, this approach ignores the fact that social perception is inherently subjective, even when labels are based on objective physical properties (Tremoulet and Feldman, 2000; Blakemore et al., 2003; Walbrin et al., 2018), and treats deviations from the intended percept as errors. Observer-based labels have been used in behavioral studies of single-agent biological motion (Davis and Gao, 2004; Johnson and Tassinari, 2005) and fMRI studies with non-social (Hebart et al., 2012) or social (Petrini et al., 2014; Nguyen et al., 2019) stimuli, although observer labels are rarely used to probe what constitutes a “Social” stimulus in the first place. Therefore, here we eschew the assumption of experimenter labels as ground truth and use observer reports to more decisively isolate brain activity associated with perceiving social interactions.

We used a large dataset ($n = 1049$ healthy adults) from the Human Connectome Project to investigate the behavioral and neural correlates of subjective (or conscious) social perception. We leveraged the following unique features of this dataset: (1) a single animation that yielded high variability in reported percepts, allowing us to isolate neural responses to conscious social perception while holding visual input constant, and (2) a task design that permitted participant reports of “Unsure” as an intermediate between “Social” and “Non-social”, allowing us to identify brain regions whose activity scaled with the extent of perceived social content. Results revealed that people show a slight behavioral bias toward perceiving information as social and that observer responses explain more variance in activity than experimenter labels in many brain regions. Occipital, temporal, and prefrontal regions showed higher responses to social information, and these differences emerged early in time and in the cortical hierarchy. Finally, internalizing traits influenced both behavior and brain activity during social signal detection.

Materials and Methods

We primarily used data from the social cognition task of the Human Connectome Project (hereafter referred to as the HCP study or HCP dataset; Van Essen et al., 2013). The dataset is openly accessible and consists of a large sample of neurotypical individuals, enabling us to study both the dominant and nondominant percepts for specific animations. The social task was one of seven cognitive tasks that were run as part of the HCP task battery (Barch et al., 2013). In this task, participants watched 10 animations, 20 s long, of which five were considered generally social and five were considered generally non-social (experimenter-assigned labels of Mental and Random, respectively). At the end of each

animation, participants indicated whether they perceived a social interaction by pressing buttons (“Social”, “Non-social”, or “Unsure”). To distinguish experimenter-assigned labels from observer responses, here we use the terms Mental and Random for the former, and “Social”, “Non-social”, and “Unsure” for the latter. In the HCP dataset, participants also completed trait-level questionnaires, which enable the study of individual differences. Here, we focused on internalizing symptoms, which include anxiety, loneliness, and social withdrawal (described later in the section titled *Correlations among traits, behavior, and neural activity*).

As participants had to wait until the end of each 20 s animation to make a response, the behavioral data in the HCP does not reveal when the perceptual decisions were made, and any differences in decision time are likely to influence the trajectory of brain activity during each trial. Hence, we additionally performed an online study on 100 neurotypical individuals [hereafter referred to as the online response time (RT) experiment] to gain insight into when decisions might have been made while watching the animation, and how decision time varied across animations and individuals.

Participants

Data from the HCP social cognition task are publicly available in the online HCP repository (<https://db.humanconnectome.org>; data for each participant, fMRI data are in the subfolders `tfMRI_SOCIAL_RL` and `tfMRI_SOCIAL_LR`; behavioral data files of the format `subID_3T_SOCIAL_runNumber_TAB.txt`, where `subID` was a six-digit number and `runNumber` was either ‘run1’ or ‘run2’). Trait scores used to study individual differences were from the restricted data category. In the demographic data reported below, the age was obtained from the restricted category and gender from the unrestricted category. We obtained complete fMRI data from 1049 individuals for the HCP social cognition task (age 22–37 years, 562 female and 486 male). Of these, 823 participants responded on all trials in a reasonable response time ($RT > 100$ ms) and were included in the behavioral data analyses. For the various fMRI analyses, depending on the comparison, participants with incomplete data were excluded. Thus, we had $n = 777$, 870, and 814 for the fMRI data analyses labelled the various fMRI analyses, namely, `RANDOM MECH`, `COAXING-BILLIARD`, and the `ALL` animations, respectively (explained in detail below under *fMRI data analysis, “Social” versus “Non-social”*). Finally, for the trait-behavior analysis, we included all participants who had complete behavior and trait data ($n = 817$), and for the trait-fMRI and trait-behavior-fMRI analyses, we included participants with behavior, fMRI, and trait data ($n = 812$).

For the online RT experiment that we conducted in July 2021, we recruited 100 neurotypical individuals (age 18–48 years, mean = 23.2, SE = 0.64) from the United States and United Kingdom via the online platform Prolific (<https://www.prolific.co>; Palan and Schitter, 2018). Before the experiment, all participants read and acknowledged the virtual consent forms in accordance with the Institutional Review Board of Dartmouth College. Participants with good-quality data ($n = 90$) were used in preliminary analyses, and of these, $n = 83$ were used to guide the `COAXING-BILLIARD` fMRI time course analyses (see below, *fMRI data analysis, fMRI time course analysis*).

Stimuli

Stimuli in the HCP study were 10 animations, each 20 s long, chosen from previous studies (Castelli et al., 2000; Wheatley et al., 2007). Longer animations had been trimmed to 20 s by the HCP researchers (Barch et al., 2013). The animations were presented in two runs with five animations each (run duration 3 min 27 s) interleaved with fixation blocks of 15 s without jitter. The order of presentation was maintained across all participants (Table 1). The number of Mental (M) and Random (R) animations were pseudo balanced within and between runs (run 1, 2M, 3R; sequence M–R–R–M–R; run 2, 3M, 2R; sequence M–M–R–M–R). Table 1 contains a list of the animations provided by the HCP and their visual properties. (Note that in this article, we drop the suffixes in the filenames -A and -B for brevity).

Each animation consisted of two or more shapes in motion (agents) with or without stationary elements (props). Seven animations (3M, 4R)

Table 1. Detailed information about the experimental stimuli used in the HCP social cognition task

Run number	Presentation sequence	Animation file names (.AVI extension)	Experimenter-assigned category	Description of agents
1. tfMRI_SOCIAL_RL	1	COAXING-B	Mental	Bigger red triangle, smaller blue triangle
	2	BILLIARDS-A	Random	Bigger red triangle, smaller blue triangle
	3	DRIFTING-A	Random	Bigger red triangle, smaller blue triangle
	4	FISHING	Mental	One circle with a fishing pole, one oblong-shaped fish
	5	RANDOM MECHANICAL	Random	Four circles, one triangle, one long rectangle of multiple colors
2. tfMRI_SOCIAL_LR	1	SCARING	Mental	Four circles, three small pink, one large blue
	2	SEDUCING-B	Mental	Bigger red triangle, smaller blue triangle
	3	STAR-A	Random	Bigger red triangle, smaller blue triangle
	4	SURPRISING-B	Mental	Bigger red triangle, smaller blue triangle
	5	TENNIS-A	Random	Bigger red triangle, smaller blue triangle

had a large red and a smaller blue triangle as agents, and the remaining three (FISHING, RANDOM MECH, and SCARING) were more diverse in the number, color, and/or form of agents and props.

For the online RT experiment, we presented the same animations used in the HCP study and in the same presentation sequence with a self-timed break after the fifth stimulus in lieu of the break between the two runs in the HCP study. In the practice phase, we randomly showed either a generally social or non-social animation (which was not one of the 10 animations used in the main task) to each participant. For a social practice example, we used MOCKING-B from the HCP repository, and for a non-social practice example, we created a two-agent animation comparable in appearance to MOCKING-B using the custom application Psyanim (https://github.com/rvarrier/HCP_socialtask_analysis/tree/main/stimuli).

The differences in physical properties that we noted above among the HCP animations could have influenced both behavior and brain activity. Hence, we factored these into our analyses either by comparing the brain activity for “Social” and “Non-social” responses within the same animation (i.e., same visual input) or by regressing out certain physical properties (i.e., the optic flow and mean brightness) before comparing individual pairs of animations in the analysis comparing time courses (see below, *fMRI time course analysis*). The presence of these visual differences also motivated our decision to perform the online RT experiment to estimate decision times and select a pair of animations with similar decision times (see below, *fMRI data analysis*). Finally, we also included animation as a grouping variable (random effect) in certain behavioral and fMRI data analyses when pooling data from multiple animations.

Experimental design

In the HCP study, participants were given the following instructions about the task: “You will now watch short clips and decide if the shapes are having a mental interaction or not. For a mental interaction, press the button under your index finger. If you are not sure, press the button under your middle finger. For a random interaction, press the button under your ring finger. After each clip, there will be a response slide. Please respond while that slide is on the screen.” They had 3 s to respond. In our online RT experiment, participants were given similar instructions but were asked to respond twice to each animation, once during the animation as soon as they made a decision (left/right arrows for “Social”/“Non-social”; referred to henceforth as the “during” phase) and a second time immediately after each animation within 3 s (left/right/down arrows for “Social”/“Non-social”/“Unsure”, similar to the HCP study; referred to henceforth as the ‘after’ phase).

Data acquisition and preprocessing

HCP social cognition task dataset

Behavioral data. In analyzing the behavioral data, we included only participants who responded to all 10 animations and in whom the response times (RT) were not unrealistically short (i.e., RTs < 100 ms were excluded), resulting in $n = 823$. Note that even if participants had arrived at a decision before the end of the video, we still need to account for the time taken to perceive the appearance of the response screen before responding (Gottsdanker, 1982).

fMRI data. The fMRI data were acquired using a 3T Skyra scanner with 2 mm isotropic voxels and a TR of 0.72 s (Barch et al., 2013, for more acquisition details). Each run comprised 274 scan volumes, and there were two runs per participant. We used minimally preprocessed voxelwise fMRI data (Glasser et al., 2013), parcellated into 268 parcels spanning the whole brain as per the Shen et al. (2013) atlas and discarded the first five scan volumes (TRs) within each run to reduce initial artifacts. Next, to make BOLD response magnitudes comparable across participants, we z-scored parcelwise time courses in each run. Further, because our analyses were to be performed at the trial level, we split the run time series into trialwise time courses of 40 s each, that is, 20 s animations (28 TRs) flanked by 10 s fixation periods (14 TRs) on either side (except for the first animation within each run, which included only six prestimulus TRs). Data preprocessed in this manner were used for all fMRI analyses except one (the time course analysis, explained later) which required comparing two individual animations, COAXING and BILLIARD. For the time course analysis, the z-normalization was done at the individual trial level to remove differences in mean activity, which were because of the order of presentation (as order was not randomized among participants). Finally, in both cases, we baseline corrected each trial time course by subtracting the signal magnitude at the trial onset (i.e., the TR immediately before stimulus onset).

Online RT experiment

In the online RT experiment, we excluded trials in which either of the two responses (during phase and after phase) were missing or when the two responses were not congruent (i.e., participants changed their response on watching the full animation). The latter was done to ensure that the response time from the during phase corresponded to the percept reported in the end to match the HCP task; however, note that the two are not perfectly comparable, since in the during phase participants did not have a choice to respond “Unsure”. Finally, as a quality check, participants with fewer than 8 of 10 good-quality (i.e., congruent) responses were also excluded, giving us 90 participants ($n = 34, 33,$ and 23 with $10/10, 9/10,$ and $8/10$ congruent responses, respectively). Based on decision times, the animation pair COAXING and BILLIARD were used in fMRI analyses to contrast “Social” and “Non-social” perception (see below, “Social” versus “Non-social” and *fMRI time course analysis*). To estimate the decision time here, we used data from 83 of the 90 participants who responded “Social” to COAXING and “Non-social” to BILLIARD congruently. The remaining seven participants either missed a response, did not give a congruent response to both animations, or did not respond to COAXING (BILLIARD) as “Social” (“Non-social”).

Data analysis

Behavioral data analysis

Using the behavioral data from the HCP, we performed four analyses to measure whether there was a general bias toward social percepts, or in other words, a shift toward “Social” responses. Our dependent variables for these analyses are described in the following paragraphs.

Table 2. Details of the three “Social” versus “Non-social” comparisons (based on observer reports) performed as part of the GLM analysis

Analysis Number	“Social” responder group	“Non-social” responder group	Rationale for the analysis	Statistical comparisons for fMRI	ID of the analysis used in figures and text
1	RANDOM MECHANICAL (Within animation, between participant)		<ul style="list-style-type: none"> Controls for low-level input Most ambiguous animation 	Two-sample <i>t</i> test	RANDOM MECH
2	COAXING (Between animation, within participant)	BILLIARD	<ul style="list-style-type: none"> Likely similar decision times during animation (as estimated from online RT experiment) 	Paired <i>t</i> test	COAXING-BILLIARD
3	All “Social” Coded 1 (“Social”) and –1 (“Non-social”), in runwise GLMs	All “Non-social”	<ul style="list-style-type: none"> Maximizes power by comparing all “Social” responses with all “Non-social” responses within participants 	One-sample <i>t</i> test after averaging runwise estimates	ALL ANIMATIONS

The first dependent variable was the percentage of “Social” and “Non-social” responses, which was estimated within participants and compared using a paired *t* test.

The second dependent variable was the decision criterion (*c*), the signal detection theory metric quantified as $\frac{-(Z(\text{HitRate}) + Z(\text{FalseAlarmRate}))}{2}$ (Stanislaw and Todorov, 1999), where *HitRate* and *FalseAlarmRate* were computed for each participant as fractions of “Social” responses for animations labeled by the experimenters as Mental and Random, respectively. Note that we do not compute other signal detection theory metrics like *d'* and bias. In this case, *d'* would be a measure of conformity to the experimenter labels (which is not of interest given our theoretical framework), and bias would be largely redundant with *c*, which already quantifies the relative magnitude of “Social” compared to “Non-social” responses. Further, *c* is preferable to bias because it is more independent of *d'* (Banks, 1970).

Third, we compared the response time (RT) differences between “Social” and “Non-social” trials. We compared RTs using both a non-parametric paired (Wilcoxon signed rank) test and a more controlled linear mixed effects (LME) analysis to further account for the differences between individual animations. The LME model (LMEM) had the form: $\log(\text{RT}) = f(\text{response}; \text{random intercepts}: \text{participant}, \text{animation})$. The factor *response* was categorical with two levels, “Non-social” (coded as the base level) and “Social”, and analysis was performed using the Python package *pymer4* (Jolly, 2018). We used the logarithm of the RT in seconds to bring the residuals of the LMEM closer to a normal distribution (which is an assumption for LMEMs).

Finally, we compared the percentage of “Unsure” responses for the two animation labels (Mental, Random) using a logistic regression as follows: $\text{uncertainty} = f(\text{stimLabel}; \text{random intercepts}: \text{participant}, \text{animation})$, where the factor *stimLabel* was categorical [Mental, Random], and the dependent variable *uncertainty* had a value of one for “Unsure” response trials and zero otherwise. Keeping Random (0) as the baseline in the analysis, positive (negative) regression coefficients for *stimLabel* would indicate a lower (higher) uncertainty in categorizing Random trials.

fMRI data analysis

GLM-based regression. Our primary approach to fMRI data analysis was a general linear model (GLM) based on animation onset and offset. We computed the regression coefficients for each animation separately for the majority of analyses. For each animation, we fitted the activity time course of each parcel to a slope regressor [line steadily increasing from zero to one from stimulus onset to offset (duration = 20 s), padded by zeros before and after] that was convolved by the Glover hemodynamic response function (HRF; Glover, 1999). (Preliminary analyses indicated that a steadily increasing slope regressor captured more variance in the BOLD data than a traditional boxcar regressor.) This renders one slope regression coefficient (β) per parcel, participant, and trial (animation). We also performed a separate GLM analysis across all animations (see below). For this analysis, we used a run-level regressor and estimated coefficients for each parcel, participant, and run. Similar to the slope regressors used at the trial level, regressor values increased (decreased) steadily during an animation labeled “Social” (“Non-social”), and were zero

at all other time points (including “Unsure” responses); thus, the run-level regression coefficient here summarizes a contrast between “Social” and “Non-social”. For each participant, we then averaged these coefficients across the two runs.

“Social” versus “Non-social”. To identify brain regions showing a consistent and generalizable difference between “Social” and “Non-social” responses, we compared the regression coefficients between “Social” and “Non-social” percepts in the following three analyses: (1) controlled for visual input, (2) controlled for decision times, and (3) across all animations (Table 2). For analyses with individual animations, we included all participants who gave a valid response to the animation(s) in that analysis, resulting in slightly different numbers of participants in each analysis (see above, *Participants*). Each analysis is described in detail below in the following paragraphs.

In the first analysis controlled for visual input, we selected the most ambiguous animation, namely RANDOM MECH, as it had the relatively most balanced “Social” and “Non-social” response groups. We excluded participants who gave an “Unsure” response to this stimulus (leaving *n* = 777) and then split regression coefficients based on observer responses (“Social”, *n* = 107; “Non-social”, *n* = 670; Fig. 1a), and compared them with two-sample *t* tests, assuming unequal variances. Although the individual groups of responders are not balanced, the actual number of individuals who responded “Social” to RANDOM MECH is still higher than in traditional fMRI studies on social perception with animations.

In the second analysis controlled for difficulty/ambiguity (COAXING vs BILLIARD), we chose two animations that were most comparable in their difficulty or ambiguity, as proxied by two measures, (1) the relative proportions of dominant and nondominant responses and (2) the time taken to arrive at a response. We used McNemar’s (1947) statistic to compare the relative proportions of dominant (“Social” and “Non-social” for COAXING and BILLIARD, respectively) and nondominant responses (“Non-social” and “Social” for COAXING and BILLIARD, respectively) in the HCP dataset (dominant, COAXING *n* = 886, BILLIARD *n* = 876; nondominant, COAXING *n* = 6, BILLIARD *n* = 16; continuity correction performed) as well as the online RT experiment (dominant, COAXING *n* = 84, BILLIARD *n* = 83; nondominant, COAXING *n* = 0, BILLIARD *n* = 1; exact binomial distribution used because of the extremely small sample sizes in the nondominant group) and found that the proportions were not significantly different in either case (see below, *Results, Decision time on whether an animation is social varies widely between individuals and animations decision time*). Response times were based on the data we obtained from the online RT experiment, where the decision time to report “Social” to COAXING (median = 3.45 s, SE = 0.27 s) and “Non-social” to BILLIARD (median = 3.7 s, SE = 0.25 s) were the closest and did not significantly differ (Fig. 2c; see below, *Results, Decision time on whether an animation is social varies widely between individuals and animations*). Hence, we compared regression coefficients for each of these two animations within participants using a paired *t* test. Note that we excluded participants who gave an uncertain or nondominant response for one or both animations (i.e., who responded to COAXING as “Non-social” or “Unsure” or BILLIARD as “Social” or “Unsure”), giving us *n* = 870 for this analysis.

In the across all animations (ALL) analysis, we also performed a more general comparison between brain activity associated with “Social” versus “Non-social” responses by identifying regions that showed a mean run-level regression coefficient that was different from zero per a one-sample *t* test (for details on how the run-wise regressor was estimated, see section, *GLM-based regression*). To minimize biases because of missed responses, we used only participants who had given all 10 responses and had complete fMRI data from both runs ($n = 814$).

Finally, we identified brain regions that were significant in all three of the above comparisons and showed changes in the same direction (either “Social” greater than “Non-social” in all three comparisons or vice versa) at a corrected threshold [false-discovery rate (FDR) $q < 0.05$, corrected for 268 comparisons (parcels)]. Hereafter, we refer to this procedure as the “intersection analysis” and the resultant parcels as “robust social perception regions.”

Comparison between using observer responses and experimenter-assigned labels. To evaluate whether observer responses actually explain more variance in the fMRI data than experimenter-assigned labels, we also ran two LMEMs for each brain parcel. The dependent variable was the trial-level regression coefficient (β) with run-normalized BOLD data (see above, *Data acquisition and preprocessing*), and both models included subject ID and animation as random effects. The fixed effect in the first model was the observer responses ($\beta = f(\text{response}; \text{random intercepts: participant, animation}); \text{response} = \{\text{“Social”}, \text{“Non-social”}\}$), and in the second model, it was the experimenter-assigned labels ($\beta = f(\text{stimLabel}; \text{random intercepts: participant, animation}); \text{stimLabel} = \{\text{Mental, Random}\}$). These models are referred to in subscripts as Obs (observer response-based) and Exp (experimenter label-based), respectively. We then assessed which label type explained more variance in the data by taking a difference between the Akaike information criteria (AIC) for each model. Lower AICs indicate better model fits, so $\text{AIC}_{\text{Obs}} < \text{AIC}_{\text{Exp}}$ indicates that the response-based model is better and vice versa. We identified parcels for which models differed in their AIC fits by at least 10, which corresponds to a relative likelihood of 99.32% for the model with the better fit (Wagenmakers and Farrell, 2004).

“Social” versus “Unsure” versus “Non-social”. We also leveraged the “Unsure” responses to identify brain regions that responded parametrically to the level of perceived socialness. We predicted that the neural response in such regions during animations ultimately marked “Unsure” would be intermediate to that of “Social” and “Non-social” responses. However, intermediate does not necessarily mean halfway, and hence we performed conjunction analyses, that is, we identified brain regions showing “Social” greater than “Unsure” and “Unsure” greater than “Non-social” (or vice versa) and took the intersection of these. We performed this analysis across all the animations using an LMEM of the form $\beta = f(\text{response, random intercept: participant})$, which was performed separately for “Social” versus “Unsure” (LMEM 1) and “Unsure” versus “Non-social” (LMEM 2). In each LMEM, *response* was a categorical variable that had the values “Social” and “Unsure” in LMEM 1 (baseline “Unsure”), and “Unsure” and “Non-social” in LMEM 2 (baseline “Non-social”). Thus, a positive LMEM estimate for *response* would indicate a higher neural response corresponding to a higher perceived socialness. From this, we identified parcels that showed the same directionality for LMEM 1 and 2 at the multiple-comparisons-corrected threshold and that were also in the set of robust social perception regions from the GLM analysis described above.

To probe whether similar parametric patterns that were seen across all animations also emerge when controlled for visual input, we again leveraged the most ambiguous animation (RANDOM MECH). We plotted the time courses for a subset of the parcels in which “Unsure” was the closest to the halfway point between “Social” and “Non-social” both in terms of the mean regression coefficient and the magnitude of activity at the end of the stimulus presentation period (20 s) for each parcel and response, the rationale being that the signal during the final time points of the animation should most closely reflect a participant’s ultimately reported percept.

fMRI time course analysis. To identify the brain regions where the earliest differences in activity between “Social” and “Non-social” percepts emerged, we performed paired *t*-tests (within participant) for each time point (TR) between BOLD responses corresponding to a pair of

“Social” and “Non-social” animations (COAXING and BILLIARD, respectively). This pair of animations was chosen because decisions on whether the animation was social or non-social were likely made at comparable times while watching them as explained earlier (see above, “Social” versus “Non-social”). To ensure that the differences in BOLD activity between COAXING and BILLIARD were not because of differences in basic visual input between the two animations, we performed these comparisons on the residual time courses obtained after regressing out two low-level visual features—total optic flow and mean brightness. We first estimated these two features for each animation frame using the *plers* software package (McNamara et al., 2017), then down-sampled the resulting time courses to match the temporal resolution of the fMRI data (i.e., the TR), *z*-normalized them, and convolved them with an HRF. We then performed a linear regression on each participant’s trial time course (including 14 TRs flanking the stimulus duration on either end, the same procedure as the slope regressors described earlier) to regress out the changes in BOLD activity related to these features. We then used the resultant residual time courses for COAXING and BILLIARD for the time course analysis. We compared these at each time point (TR) and for each parcel using paired *t* tests (within participant). For each parcel, we thus identified the earliest time point at which BOLD activity begins to diverge (i.e., $p < 0.05$). As additional consistency checks, we (1) only performed this analysis in the robust social perception regions from the GLM intersection analysis and (2) selected a TR *t* as the divergence point only if the difference between “Social” and “Non-social” timecourses at $t + 1$ was also significantly different ($p < 0.05$) in the same direction.

Note that this analysis does not factor in the hemodynamic lag. This is because although the HRF peaks a few seconds after an event (in our case, the animation onset), the neural responses—and corresponding start of the BOLD response—should have begun nearly instantly on stimulus presentation (Friston et al., 1994), so here we investigated where these earliest changes could be observed. Further, in using the median decision times from the online RT experiment for COAXING and BILLIARD as the expected decision time for the HCP dataset, we did not factor in the motor response delay (i.e., time taken after a decision has been made to press a button) in the online RT experiment. Hence, it is possible that some of the predecisional processes closer to the decision time may have in fact been postdecisional. Although we cannot exclude this possibility, this was unlikely because motor responses on arriving at a decision are typically quicker than the TR used in the HCP task (0.72 s).

We also did not multiple-comparisons correct across time points in this analysis as the primary goal was to identify the earliest differences in activity, and to infer this correctly, false negatives are undesirable. Still, in identifying the earliest time points, we only selected a region if the subsequent time point was also significant ($p < 0.05$ uncorrected), limiting the odds of a false positive further by 95%.

We also did not perform this analysis within the same animation (RANDOM MECH) and across all animations like in the GLM analysis (see above, “Social” versus “Non-social”) because of the heterogeneity in decision times for RANDOM MECH and across animations (Fig. 2c). This heterogeneity means that the neural processes at each time point could have been vastly different between individuals and animations, thus clouding any potential interpretations from these comparisons.

Correlations among traits, behavior, and neural activity

Past work has shown that within the neurotypical population, social perception covaries with traits like loneliness, anxiety, psychopathy, and autism-like phenotypes (Sacco et al., 2016; Lessard and Juvonen, 2018; Desai et al., 2019; Abassi and Papeo, 2022; Lisøy et al., 2022; Williams and Chakrabarti, 2022). In particular, individuals high on internalizing traits such as loneliness and anxiety are more sensitive to social cues (Gardner et al., 2005), tend to form illusory social connections by anthropomorphizing inanimate objects (Epley et al., 2008; Powers et al., 2014), and show smaller gray matter volumes in a brain region typically associated with social processing, the posterior STS (pSTS; Kanai et al., 2012). Here, we probed whether internalizing traits affect behavior and/or brain activity associated with social perception using the internalizing *T* score provided by the HCP (Barch et al., 2013). This score is based on

participants' responses to the internalizing dimension questions from the Achenbach Adult Self-Report questionnaire (ASR; Achenbach et al., 2017). Internalizing symptoms refer to symptoms like anxiety, depression, and withdrawal, and are typically contrasted with externalizing behaviors such as rule breaking and aggression. The ASR was designed to assess behavioral, emotional, and social functioning across a wide spectrum of the population, so it is sensitive to individual differences (i.e., produces a range of scores) even in healthy/subclinical populations. Specifically, we used the participant-level internalizing T score (labeled ASR_Intn_T in the HCP dataset), which is normalized for age and sex (mean = 48.72, STD = 10.75, range = 30–97; see Fig. 7a–c for full distribution) and sums across the three ASR scales Anxious/Depressed, Withdrawn, and Somatic Complaints.

To assess whether the internalizing score relates to a behavioral bias toward “Social” percepts, we compared participants' internalizing scores with the following behavioral variables: (1) the difference between the percentage of “Social” and the percentage of “Non-social” responses (calculated as percentages to control for missing data; Spearman's rank correlation, r_s); (2) responses to the most ambiguous animation, RANDOM MECH, specifically comparing “Non-social” to “Social” or “Unsure” responders (two-sample t test); and (3) the percentage of “Unsure” responses on Mental versus Random trials (Spearman's correlation). We tested the specificity of these correlations by additionally performing correlations with externalizing scores (ASR_Extn_T in the HCP dataset) and comparing the strength of the relationships using the CorrelationStats package (<https://github.com/psinger/CorrelationStats>). For RANDOM MECH, we used the LMEM $traitScore \sim f(respGroup, traitType, respGroup * traitType \text{ interaction}, random \text{ intercept: participant})$ to study how the internalizing and externalizing trait scores [dependent variable $traitScore$; factor $traitType$ with two levels (internalizing, externalizing)] vary for each response group [factor $respGroup$ with two levels (“Non-social”, “Social”/“Unsure”)] and with $participant$ as a random effect.

To quantify whether and where internalizing traits relate to brain activity while scanning animations for social information, for each parcel we performed an LME analysis where the dependent variable was the trial-level slope regression coefficient, the fixed factor was internalizing score, and the random factor was animation. This yields brain regions that respond proportionately to internalizing score in that individual across animations and parcels.

Finally, to check for interactions among social percepts, internalizing symptoms, and neural activity, we tested how much the difference between neural responses to “Social” and “Non-social” trials depended on internalizing symptoms. For this, we fitted regression coefficients from the GLM analysis across all animations (ALL), which represents the contrast between activity to “Social” versus “Non-social” responses (one estimate per participant; see above, *GLM-based regression and “Social” versus “Non-social”*), to the internalizing symptom scores (also one estimate per participant) in a linear regression.

Data availability

All the code for analyzing data from both the HCP and online RT experiment, as well as the anonymized data from the online RT experiment is available at https://github.com/rvarrier/HCP_socialtask_analysis.

Results

In this study, we used behavioral and fMRI data from the HCP social cognition task to characterize the behavioral and neural processes underlying conscious perception of social interactions. We started by evaluating the behavioral data for any response bias, that is, are people more inclined to declare information “Social” (as opposed to “Non-social”)? We next identified brain regions that robustly differentiated between “Social” and “Non-social” percepts, including a subset that showed a parametric response pattern to degrees of perceived socialness, and showed that observer responses explain more variance than experimenter-assigned labels in activity levels of many regions. Next, we used a time-point-by-time-point analysis to identify

where and when brain activity begins to diverge according to whether social information is subjectively perceived. Finally, we studied the relationship between internalizing behavior scores, tendency toward social percepts, and brain activity while scanning for social information. As a reminder, here, we use the terms “Mental” and “Random” to refer to experimenter-assigned stimulus labels, and “Social” and “Non-social” to refer to observers' actual reported percepts of those stimuli.

Some animations are more ambiguous than others

First, we examined the degree to which participants' percepts of “Social” versus “Non-social” information agreed with one another as well as the intended stimulus category. In the HCP social cognition task, participants passively watched 10 animations, 20 s long, of geometric shapes (Heider–Simmel-like; Castelli et al., 2000; see above, *Materials and Methods, Stimuli*) and then made a behavioral response—“Social”, “Non-social”, or “Unsure”—to indicate whether they perceived a social interaction in the animation. Five animations were intended to evoke social interactions (experimenter-assigned Mental) and five were not (experimenter-assigned Random). Although on average, participants' percepts aligned with experimenter labels, the degree to which animations were perceived as “Social” and “Non-social” varied considerably. This was true in both the HCP behavioral data and the secondary online dataset (online RT experiment) we collected to study the time taken for individuals to arrive at decisions while watching each animation (Figs. 1a, 2a). Although animations like DRIFTING and BILLIARD were seen almost unanimously as “Non-social”, animations like RANDOM MECH and FISHING had a higher percentage of the nondominant percept as well as “Unsure” responses. This underscores the need to use participants' own percepts to categorize what is or is not “Social” rather than experimenter-assigned labels. In later analyses, we leveraged this ambiguity by comparing neural activity corresponding to “Social” and “Non-social” responses within the most variably perceived animation (RANDOM MECH), thereby isolating activity associated with a conscious social percept while controlling for visual input.

Responses are biased toward “Social”

Next, we used behaviorally reported percepts to determine whether there was a response bias toward “Social”. We hypothesized that evolutionarily, there may be a bias toward perceiving information as social because the cost of a false positive (i.e., mistakenly thinking someone is trying to engage you in a social interaction) is less than that of a false negative (i.e., missing out on social cues that are important for group dynamics, reproduction, and survival). We predicted that this bias would manifest as a higher “Social” response rate, shorter response times for “Social” percepts, and more “Unsure” responses to animations labeled Random by experimenters (because of a reluctance to declare something entirely non-social). Our findings are described next.

“Social” responses are more frequent

Comparing the frequency of “Social” and “Non-social” percepts for each participant (limited to trials where participants were sure of their response, that is, excluding “Unsure” trials) showed that the percentage of “Social” responses was higher (mean = 52.9%, SE = 0.29%) than that of “Non-social” responses (mean = 47.1%, SE = 0.29%; paired t test, $t = 9.96$, $p < 10^{-21}$; Fig. 1b).

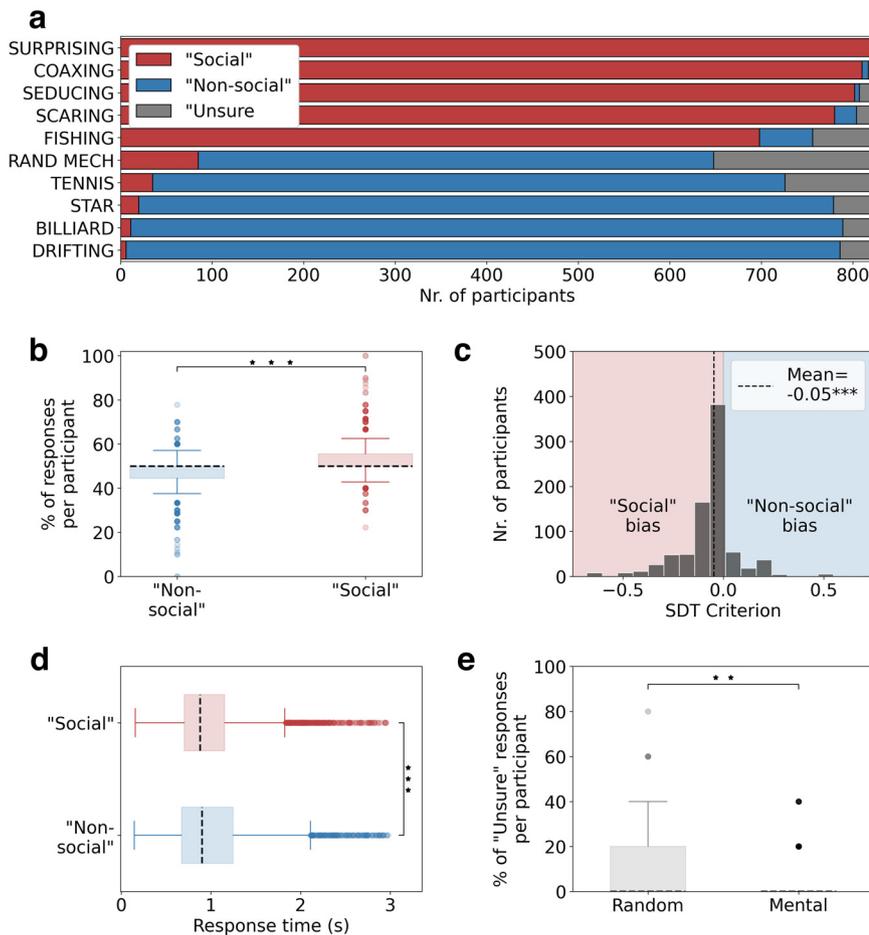


Figure 1. Behavioral data from the HCP participants ($n = 823$) show a bias toward “Social” responses. **a**, Number of responses per type (“Social”, “Unsure”, “Non-social”) for each animation (sorted from most to least “Social”). **b**, Percentages of “Social” and “Non-social” responses. “Social” responses were more frequent ($t = 9.96$, $p < 10^{-21}$, paired t test). **c**, Signal detection theory metric criterion c across participants based on experimenter-assigned labels. Mean criterion was negative ($\bar{c} = -0.05$, Wilcoxon signed-rank test statistic, 26813; $p < 10^{-17}$), indicating a bias toward false alarms (i.e., declaring an animation labeled Random by experimenters as “Social”). **d**, Response time for “Social” and “Non-social” responses. “Social” responses tended to be quicker (Wilcoxon signed-rank test statistic, 144885; $p < 10^{-3}$). **e**, “Unsure” responses for animations labeled Mental and Random by experimenters. There was a higher percentage of “Unsure” responses for Random animations (LMEM, Est. = -2.15 , $p < .005$; $**p < 0.001$, $***p < 0.0001$).

The response criterion further shows a bias toward “Social”

Next, we computed criterion (c), a metric from signal detection theory that quantifies response biases. If the mean c is significantly different from zero, this suggests a bias in responses toward “Social” ($\bar{c} < 0$) or “Non-social” ($\bar{c} > 0$). We found that criterion was significantly negative at the population level (mean = -0.05 , SE = 0.01; Wilcoxon test, test statistic = 26813, $p < 10^{-17}$; Fig. 1c), further confirming the response bias toward “Social”. In this computation, we used the experimenter-assigned labels to show that although the experimenters aimed to create a balanced set of five Mental and five Random animations, actual observer reports indicate that individuals ended up perceiving more animations as “Social”. Thus, percepts did not fully conform to the expectations of the experimenters.

Responders may have been quicker to declare something as “Social” than “Non-social”

Next, to get at a more subconscious measure of perceptual decision-making for social information, we compared response times between “Social” (mean = 0.87 s, SE = 0.009 s) and “Non-social”

(mean = 0.9 s, SE = 0.012 s) responses (Fig. 1d) and found that “Social” responses were overall faster (Wilcoxon test, test statistic = 144885, $p < 10^{-3}$). As response times could differ by animation because of their heterogeneity, we additionally performed an LME analysis with *response* [“Social” or “Non-social” (baseline)] as the fixed effect, and both *animation* and *participant* as random effects. We observed a trend toward shorter RTs for “Social” responses, but this did not reach significance (LMEM estimate = -0.04 , $p = 0.1$).

“Unsure” responses were more common for animations intended as Random compared to those intended as Mental

We compared the distribution of “Unsure” responses between animations that were intended to be “Social” (Mental) or “Non-social” (Random) and noted a higher percentage of “Unsure” responses in the animations intended as Random (mean = 9.4%, SE = 0.5%; Fig. 1e) compared to those intended as Mental (mean = 2.7%, SE = 0.3%). This indicated that people were more hesitant to label something “Non-social” (as opposed to “Social”) when their confidence is low. In other words, they err on the side of false alarms rather than misses; this fits with the idea that misses are likely costlier than false alarms. We formally compared the frequency of “Unsure” responses using log-istic regression with Mental (coded 1) and Random (coded 0) labels as two levels of the fixed effect term *stimType*, and *participant* and *animation* as random effects. Results showed higher uncertainty on Random trials even after accounting for the differences in animations (Est. = -1.61 , $p = 0.005$).

To summarize, the behavioral data overall showed a bias toward “Social” responses based on frequency of each response type, response times, and degree of uncertainty.

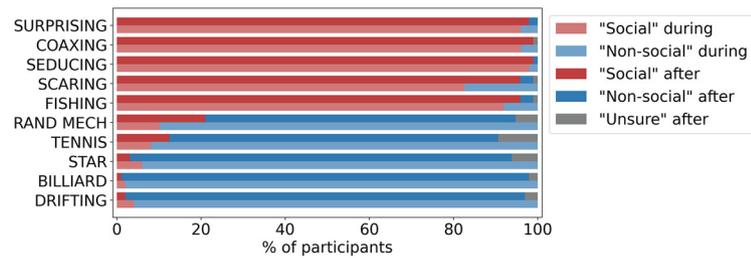
Decision time on whether an animation is social varies widely between individuals and animations

In the HCP study, participants had to wait until the end of each animation (lasting 20 s) to make a behavioral response. However, the decision on whether an animation was “Social” or “Non-social” was presumably made sometime during passive viewing, although the decision time could have varied widely across animations and participants. This variability, in turn, might influence the time course of brain activity (e.g., visual attention for the same animation may be different when a participant makes a decision 2 s into the animation vs 15 s into it). Hence, getting information on when decisions could likely have been made during each animation was critical to modeling and interpreting neuroimaging data. To this end, we performed an independent online behavioral study using the same animations where participants (final $n = 90$) were instructed to indicate

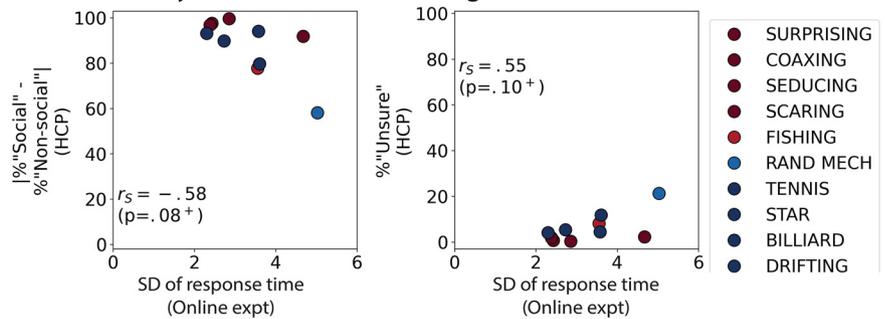
their percepts as soon as they had arrived at a decision (the during phase). To compare the results with the HCP study, participants were also instructed to respond at the end of each trial (the after phase).

The consensus across participants of which animations were generally “Social” versus “Non-social” in the online sample was comparable to that of the HCP sample (compare Fig. 2*a* with Fig. 1*a*). As a corollary to this, the animations with high variability in decision times in the online RT experiment also tended to have less consensus across participants in the HCP study, the latter operationalized as (1) the lower absolute value of the difference between the percentage of “Social” and percentage of “Non-social” animations (Fig. 2*b*, left) and (2) the higher number of “Unsure” responses (Fig. 2*b*, right). The reaction time data from the during phase (Fig. 2*c*) showed that although most responses were made in the earlier half of the 20 s animations, there was a high variability in decision time both within and across animations. This means that the brain activity corresponding to an especially ambiguous animation (e.g., SCARING, RANDOM MECH) could have been vastly different even among participants who reported the same percept for these, depending on when each participant made their decision and how this affected their attention before and after the decision. Hence, we identified two animations with the most comparable decision times, namely, COAXING (median = 3.45 s, SE = 0.27 s), a predominantly “Social” animation, and BILLIARD (median = 3.7 s, SE = 0.25 s), a predominantly “Non-social” animation, whose decision times were not significantly different (Wilcoxon signed-rank test paired, $t = 1619$, $p = 0.57$). These animations also did not differ (McNemar test statistic = 0, $p = 1$, exact correction done) in their proportion of dominant (“Social” and “Non-social” for COAXING and BILLIARD, respectively) and nondominant responses (“Non-social” and “Social” for COAXING and BILLIARD, respectively; see above, *Materials and Methods*, “Social” versus “Non-social”). This was the case with the HCP dataset too (McNemar test statistic = 3.7, $p = 0.055$, continuity correction done), although the differences in proportions was close to significance despite a large proportion of the participants ($n = 870$) reporting the dominant percepts to COAXING and BILLIARD. Therefore, we used this pair of animations in later analyses as a control for stimulus difficulty/ambiguity.

a Responses during and after each animation



b Uncertainty in decision time during an animation



c Response time during each animation

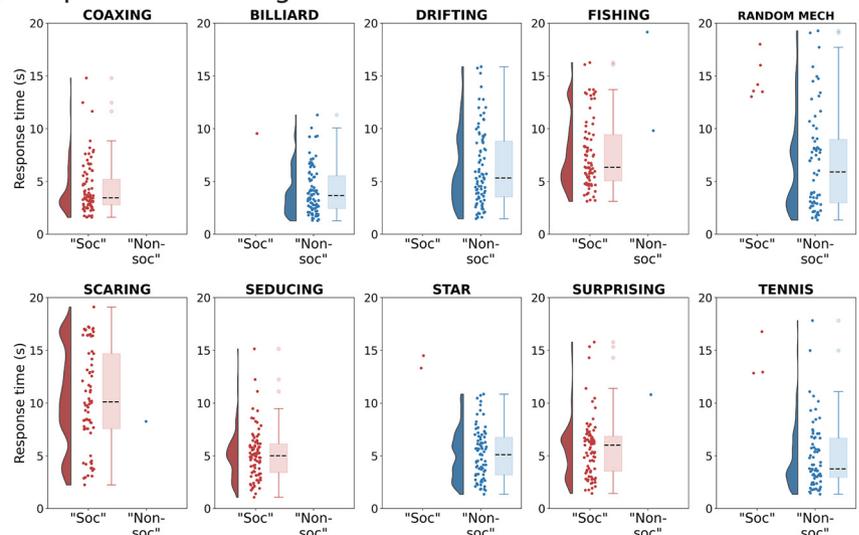


Figure 2. Results of the online RT experiment to characterize decision time for each animation. *a*, Number of “Social”, “Non-social”, and “Unsure” responses per animation made during (lighter shades) and after (darker shades) each animation. Order of animations on the y-axis is the same as for the HCP data in Figure 1*a*. The degree to which animations were reported “Social” is comparable to the HCP behavioral data in Figure 1*a*. *b*, SD of response time while watching each animation (in seconds, x-axis) versus two indicators of uncertainty from the HCP behavioral data on the y-axes (left, absolute difference between number of “Social” and “Non-social” responses, an indicator of how definitive responses for this animation were across participants; right, percentage of “Unsure” responses). Spearman (rank) correlation shows a trend ($p \leq 0.1$, marked +) for the animations with higher variation in response times in the online RT experiment x-axes to also have a less definitive response (left) and a higher percentage of “Unsure” responses (right) in the HCP behavioral data. Dots are colored according to each animation’s average perceived socialness (average response from *a*). *c*, Distribution of response times for “Social” and “Non-social” responses while watching each animation (in seconds). As seen in *b*, decision times varied more for some animations than others. Note that COAXING (dominantly “Social”) and BILLIARD (dominantly “Non-social”) had similar mean decision times that were both relatively early in the animation.

Much of the brain responds more strongly to what is perceived as social information

In the next set of analyses spanning this and the next two sections, we used the HCP fMRI data to understand where and when the brain distinguishes social from non-social information. For all fMRI analyses, whole-brain data were parcellated into 268 regions covering the cortex, subcortex,

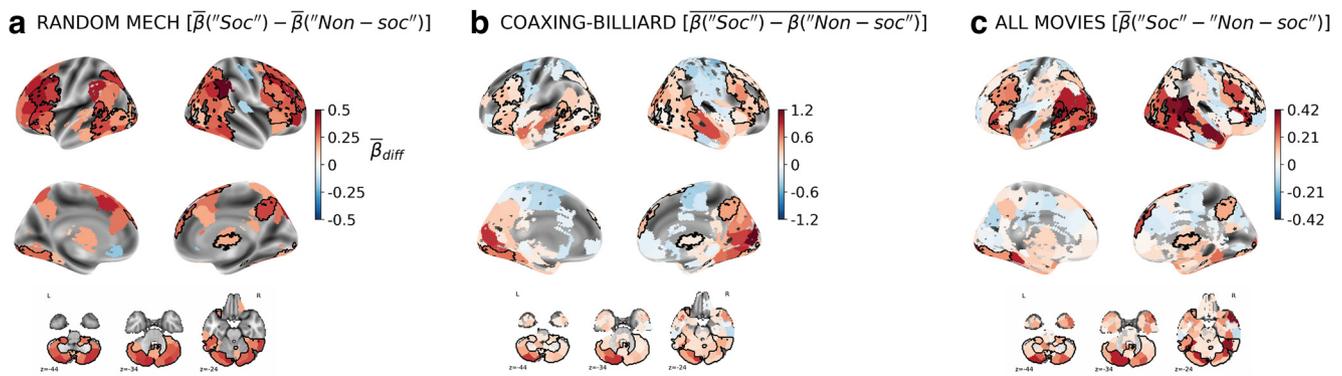


Figure 3. Regions showing differential activity between “Social” and “Non-social” percepts. **a–c**, Mean differences between GLM regression coefficients (β) for (**a**) RANDOM MECH [mean (RANDOM MECH “Social”) – mean (RANDOM MECH “Non-social”), (**b**) COAXING–BILLIARD [mean (COAXING “Social”) – BILLIARD “Non-social”), and (**c**) ALL (estimated from run-level regressors, see *Materials and Methods*). Colored regions are significant at an uncorrected threshold ($p < 0.05$) in each of the three analyses, whereas black contours in **a–c** show the robust social perception regions significant after correction for multiple comparisons (FDR $q < 0.05$) in all three analyses. Color bar ranges are different among the three subplots as each was estimated separately using different analyses, and, hence, the values should not be directly compared.

Table 3. List of the robust social perception regions (parcels) identified from the intersection analysis (compare black contours in Fig. 3), together with their Montreal Neurological Institute (MNI) coordinates

Lobe	Shen parcels (MNI [x, y, z] coordinates)
Occipital	<ul style="list-style-type: none"> • Bilateral lateral occipital cortex ([41, –75, 28], [30, –83, 20], [45, –74, 3], [–32, –87, 13], [–43, –70, –14], [–48, –67, 1], [–36, –84, –4]). • Bilateral Occipital pole ([31, –92, –11], [–22, –97, –10]). • Bilateral occipital fusiform gyrus ([37, –69, –17], [18, –83, –11], [–26, –63, –12], [–15, –84, –13]).
Temporo–occipital	<ul style="list-style-type: none"> • Bilateral middle and inferior temporal gyrus ([47, –60, –15], [55, –56, –5], [61, –43, –18], [42, –46, –23], [–60, –50, –14], [–47, –40, –24], [–43, –52, –17]).
Parietal	<ul style="list-style-type: none"> • Bilateral posterior superior temporal gyrus ([49, –58, 14], [59, –44, 9], [–58, –46, 6]). • Right precuneus ([6, –57, 38]).
Frontal	<ul style="list-style-type: none"> • Bilateral occipito-parietal ([32, –61, 49], [48, –62, 35], [–28, –62, 40]). • Bilateral middle and inferior frontal gyrus ([41, 15, 48], [54, 25, 1], [48, 36, 15], [40, 18, 29], [–46, 28, 27], [–39, 17, 47], [–53, 18, 11]). • Bilateral lateral precentral gyrus ([40, 4, 34], [–46, 8, 29]). • Bilateral frontal operculum ([37, 21, 6], [–32, 22, 6]). • Bilateral frontal pole ([29, 51, 19], [9, 53, 24], [–10, 56, 30]). • Bilateral anterior insula ([37, 21, 6], [–32, 20, –16], [–32, 22, 6]). • Right superior frontal gyrus ([15, 37, 49], [24, 31, 36], [14, 6, 65], [25, 12, 49]). • Left frontal orbital cortex ([–46, 28, –7], [–32, 20, –16]).
Sub–cortical	<ul style="list-style-type: none"> • Bilateral cerebellum (large parts of it) ([32, –78, –40], [39, –75, –30], [23, –36, –43], [7, –68, –38], [46, –46, –43], [16, –47, –53], [12, –84, –35], [20, –73, –50], [37, –57, –33], [23, –72, –29], [42, –64, –49], [7, –54, –34], [30, –36, –31], [–6, –66, –38], [–21, –70, –49], [–40, –75, –29], [–30, –80, –40], [–43, –64, –46], [–10, –82, –32], [–46, –47, –43], [–26, –70, –31], [–24, –38, –44]). • Right thalamus ([6, –10, 5]).

MNI, Montreal Neurological Institute.

and cerebellum using the Shen et al. (2013) atlas to ease the computational burden of voxelwise analyses.

In the first fMRI analysis, we focused on the question of where by comparing overall neural responsiveness while viewing animations ultimately deemed “Social” versus “Non-social”. In addition to regions along the STS that are known to be involved in animacy and interaction perception, we hypothesized that differences might emerge as early as visual regions. We compared “Social” and “Non-social” responses using a GLM approach, again using the participant’s reported percept rather than the experimenter-assigned label as input to the model, in three separate contrasts to ensure results were robust to different confounding factors: (1) within the single most ambiguous animation (RANDOM MECH), which controls for visual input (as all participants saw the same animation, but reported different percepts; across participants); (2) between two animations with similar decision times (COAXING versus BILLIARD),

to control for the effect of when the decision was likely made on the time course of brain activity during passive viewing (within participants); and (3) across all 10 animations, to maximize power and ensure generalizability (within participants). We then took the intersection of the regions showing a significant difference in all three analyses.

In total, 70 parcels showed a higher activity for “Social” compared to “Non-social” (FDR $q < 0.05$, black contours in Fig. 3, Table 3) consistently across all three comparisons. (No parcel showed a higher activity for “Non-social” compared to “Social” across analyses, although there were some results in this direction in the uncorrected analyses; Fig. 3b–c). Of these, 66 parcels showed positive (i.e., above baseline) activations for both the “Social” ($\beta_{\text{“Social”}} > 0$) and “Non-social” ($\beta_{\text{“Non-social”}} > 0$) responses for both RANDOM MECH and COAXING–BILLIARD, suggesting that on the whole, much of the brain showed higher activation and not lower deactivation to “Social” compared with “Non-social”.

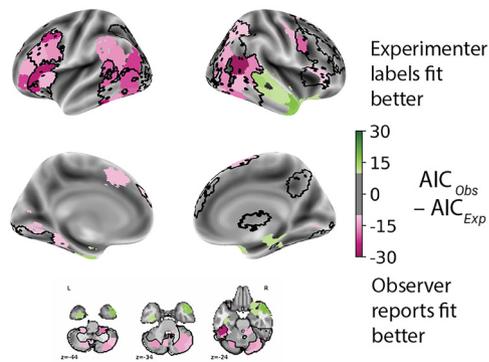


Figure 4. Models based on observer reports, compared with experimenter labels, better explain brain activity in robust social perception regions. Comparison between two linear mixed-effects models fitting trial-level GLM regression coefficients (β) as a function of either (1) observer responses (“Social” vs “Non-social”) or (2) experimenter labels (Mental vs Random). Parcels are colored by the difference in AIC between the two models and were thresholded at 10 (i.e., parcels with $|AIC| < 10$ are not plotted); parcels colored pink (difference in $AIC < -10$) indicate a better fit for the observer response-based model, and those colored green (difference in $AIC > 10$) indicate a better fit for the experimenter label-based model. Both models included *participant* and *animation* as random effects. Black contours correspond to the robust social perception regions identified in Figure 3, and these largely overlap with parcels that show better fits for the observer-based models (pink).

These parcels spanned the occipitotemporal cortex, the prefrontal cortex, the cerebellum, and some subcortical regions (Table 3). Hereafter, we refer to this set of 70 parcels as the “robust social perception regions” as they show both specific (after controlling for visual input and decision time) and generalizable activation associated with the subjective experience of a social percept.

Subjective percepts better explain brain activity in robust social perception regions

Next, we tested whether observer response-based labels (“Social”, “Non-social”) explained more variance in the neural data than experimenter-assigned labels (Mental, Random) by comparing models based on each label type. Overall, across all 268 parcels, response labels better explained brain activity ($AIC_{Obs-Exp}$, mean = -2.23 , SE = 0.47 ; one-sample t test, $t = -4.77$, $p < 10^{-5}$). Of these, 44 parcels were better fit by the observer-based model ($AIC_{Obs-Exp} > 10$) and 11 parcels were better fit by the experimenter-based model ($AIC_{Obs-Exp} < -10$). Observer-based labels better fit the neural activity in bilateral occipitotemporal regions, left prefrontal cortex, and the cerebellum (Fig. 4, pink-colored parcels), several of them overlapping with the robust social perception regions identified in the GLM analysis above (Fig. 4, black contours), whereas experimenter-assigned labels better fit the neural activity in the right temporal cortex (Fig. 4, green-colored parcels). Overall, this suggests that activity in the robust social perception regions reflect the conscious perception of social information rather than merely incoming visual input.

Some brain regions show parametric responses to degree of perceived socialness

The previous analysis identified social-information-processing regions that robustly showed a higher response to information ultimately reported as “Social”. By leveraging “Unsure” responses as an intermediate level of perceived socialness between “Social” and “Non-social”, we further probed the neural correlates of conscious social perception; that is, an “Unsure” response would indicate that some evidence for a social interaction was detected but not enough to be fully confident in a “Social” response.

Specifically, we identified regions showing parametric responses, that is, $\beta_{\text{“Social”}} > \beta_{\text{“Unsure”}} > \beta_{\text{“Non-social”}}$ [condition “Social” (S) > “Unsure” (U) > “Non-social” (NS)] or $\beta_{\text{“Social”}} < \beta_{\text{“Unsure”}} < \beta_{\text{“Non-social”}}$ (condition $S < U < NS$) using conjunction analyses across all animations (see, *Materials and Methods*). We further limited this analysis to the robust social perception regions (black contours in Fig. 3).

Several parcels showed a consistent $S > U > NS$ response pattern (Fig. 5a–c). These were located in posterior and inferior parts of the temporal cortex including parts of the motion-processing region V5/MT (with more parcels in the right hemisphere), middle and inferior frontal gyrus, precuneus, right thalamus, and postero-lateral parts of the cerebellum. Other regions that showed a differential response to “Social” compared with “Non-social” but do not show up here, such as the superior temporal and occipital regions, posterior parietal regions, and superior frontal regions, likely have more dichotomous responses to any amount of social content [(“Social”, “Unsure”) > “Non-social”] or only a high level of evidence in favor of “Social” [“Social” > (“Unsure”, “Non-social”)].

To verify that similar parametric patterns emerge when controlling for visual input, we plotted the time courses for each response type for a subset of the parcels showing parametric responses pattern to the most ambiguous animation (RANDOM MECH; Fig. 5d; see above, *Materials and Methods* for how these parcels were chosen). Visualizing these time courses confirmed that these regions show parametric neural responses to degrees of reported socialness, albeit with large error bars for the smaller groups (“Social” and “Unsure”).

Thus, it appears that many regions, predominantly in temporal, occipital, and subcortical areas, show a graded response to degree of social information. This result further underscores how using observer-based labels can increase sensitivity and specificity in linking brain activity to conscious experience.

Processing of social versus non-social information diverges early in time and in the cortical hierarchy

The previous analyses showed that several regions spanning the whole brain are more responsive to information that is ultimately reported as social (vs non-social). However, given that these analyses modeled the entire 20 s animations, any differences, especially in early visual regions, could reflect (1) the accumulation of evidence that led to the perception of an animation as “Social”, (2) the consequence of having perceived an animation as “Social” (i.e., top-down attention effects on sensory regions), or (3) a combination of both. To gain a better understanding of the dynamics of evidence accumulation leading to a “Social” percept, we compared BOLD activity at each time point (TR) after stimulus onset to determine the time point of earliest divergence between “Social” and “Non-social” percepts.

To ensure that the differences observed at each time point are comparable in terms of the underlying cognitive processes (i.e., evidence accumulation vs decision-making versus postdecisional processes), we performed this analysis on the animation pair that had the most comparable decision times in our auxiliary behavioral experiment, namely COAXING and BILLIARD. Decision times for these animations were both early and close in time (see above, *Materials and Methods*, “Social” versus “Non-social” and *Results*, *Decision time on whether an animation is social varies widely between individuals and animations*; Fig. 2c). These animations were similar visually with the same two triangular agents on the screen (Table 1); nevertheless, they did vary in their temporal dynamics and some low-level visual features. To minimize the effect of these on the BOLD activity, we regressed two low-

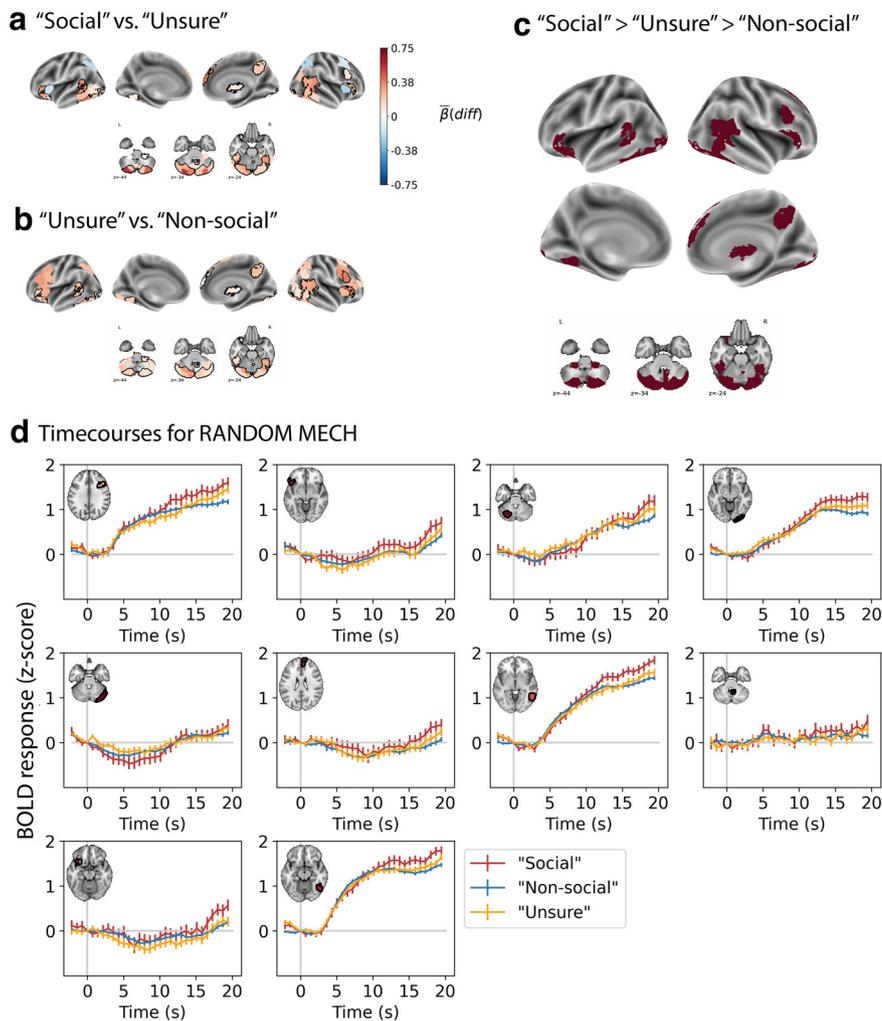


Figure 5. Brain regions showing parametric responses to social content. **a, b**, Colored parcels show mean differences (FDR $q < 0.05$) in slope regression coefficients (“Social”–“Unsure” and “Unsure”–“Non-social”) within the robust social perception regions (compare black contours in Fig. 3). **c**, Black contours and the dark red regions in **c** highlight the 35 parcels that showed a graded response to perceived socialness (“Social” greater than “Unsure” greater than “Non-social” or vice versa; in other words, the intersection of **a** and **b**). **d**, Time courses for the most ambiguous animation (RANDOM MECH) in a subset of 10 of the parcels from **a**, confirming that activity associated with “Unsure” percepts is intermediate to “Non-social” percepts even when controlling for visual input.

level visual features (total optic flow and mean brightness) from the BOLD responses of each animation and participant and compared the residual COAXING and BILLIARD time courses at each TR. To guard against spurious fluctuations early in the animations, we again limited our analysis to the robust social perception regions (compare Fig. 3, black contours).

In many regions, differences in brain activity between “Social” and “Non-social” percepts emerged early, that is, in TRs 1–3 after stimulus onset (Fig. 6a). These early differences were seen in both hemispheres in posterior regions such as the fusiform gyrus, lateral occipital cortex, pSTS; posterior parts of the cerebellum; and in frontal areas such as the lateral precentral gyrus, posterior parts of the middle and inferior frontal gyrus (IFG), the orbitofrontal cortex in the left hemisphere, and the IFG and supplementary motor area in the right hemisphere. Later TRs, which are more likely to reflect postdecisional activity, showed divergences in the bilateral inferior and superior frontal regions, the right precuneus, bilateral intraparietal sulcus, and bilateral posterior cerebellum.

To visualize the earliest differences in the posterior regions and to understand how generalizable these dynamics are, we plotted (Fig. 6b) the residual BOLD time courses for COAXING–BILLIARD (left column, our main analysis) alongside the averaged “Social” and “Non-social” time courses across all the other animations (all except COAXING–BILLIARD, middle column) and within the most ambiguous animation (RANDOM MECH, right column). The two latter analyses are not as well suited to pinpointing when differences emerged because decision times were likely more variable across individuals and animations (per our online RT experiment), thus making time courses noisier and less comparable. Despite this, we see similar relative trends in these posterior regions (each row in Fig. 6b) on when and how they distinguish between “Social” and “Non-social” reports. Responses emerged much later for the all except COAXING–BILLIARD condition in line with the later and more variable decision times for most animations; Fig. 2c). When comparing within the same animation (RANDOM MECH), we see trends emerging early, although the magnitudes are smaller, and the errors for the “Social” responder group are large, possibly because of the smaller group size ($n = 107$) compared with the majority percept of “Non-social” ($n = 670$). Note that the latter two time courses are plotted only for visual examination and that we did not perform statistical analyses here. Also note that the directionality of the difference between “Social” and “Non-social” should not be strongly interpreted especially in the case of COAXING–BILLIARD as despite our attempts to normalize activity at the trial level (see, *Materials and Methods, GLM-based regression*), order effects (COAXING was always the first stimulus in the first run, immediately followed by BILLIARD), and/or the shape of the hemodynamic response (i.e., presence of initial dip) could have affected the BOLD response between trials.

To summarize, while watching an animation that was eventually reported as “Social”, differences in brain activity emerged early across much of the brain, involving both ventral visual processing regions and occipitotemporal regions involved in action and animacy detection as well as social cognition. The early reactivity in these regions is in line with the recently suggested third visual pathway, which projects directly from early visual cortex to the superior temporal sulcus and is specialized for social perception (Pitcher and Ungerleider, 2021).

Individual differences in behavior and brain activity while viewing animations covary with internalizing symptoms

Finally, we explored whether individual differences in behavioral and neural responses to social animations covaried with trait-level

measures. Specifically, we focused on internalizing symptoms from the Achenbach Adult Self-Report Scale, because past work has shown that certain internalizing traits (e.g., loneliness, anxiety) are associated with a stronger tendency to perceive visual cues as socially salient. We hypothesized that individuals with higher internalizing scores would show stronger behavioral and neural reactivity to potentially social information.

Using the behavioral data, we tested whether the response bias toward “Social” (compare Fig. 1*a*) was even stronger for individuals higher on internalizing symptoms. Indeed, there was a positive relationship between the bias toward “Social” responses and internalizing score (Spearman’s rank correlation $r_s = 0.10$, $p = 0.003$, Fig. 7*a*). We tested the specificity of this relationship by contrasting it to the correlation with externalizing trait scores, which index more acting out behaviors like rule breaking and aggression, and have not been linked to social perception tendencies (though note that internalizing and externalizing symptoms were correlated, $r_s = 0.51$, $p < 10^{-55}$). The correlation with externalizing symptoms was not significant ($r_s = 0.06$, $p = 0.096$), although the two correlations were not significantly different ($t = 1.3$, $p = 0.094$). Furthermore, individuals with higher internalizing scores were more likely to give a “Social” or “Unsure” (as opposed to “Non-social”) response to the most ambiguous animation, RANDOM MECH (“Social” or “Unsure”, mean = 49.3, SE = 0.69; “Non-social”, mean = 47.7, SE = 0.45; unpaired t test, $t = 2.05$, $p = 0.04$; Fig. 7*b*). Mean externalizing symptoms were also higher for the “Social” or “Unsure” group (mean = 49.3, SE = 0.57) compared with the “Non-social” group (mean = 47.9, SE = 0.38). Although the difference in externalizing symptoms was smaller (unpaired t test, $t = 1.95$, $p = 0.051$), it was not significantly different from the internalizing symptoms [interaction between response and score type (internalizing vs externalizing), $p = 0.63$].

Finally, individuals with higher internalizing scores were also more likely to give an “Unsure” response to animations intended as Random ($r_s = 0.098$, $p = 0.005$) but not to animations intended as Mental ($r_s = -0.024$, $p = 0.49$), indicating a preference for false alarms over misses when it comes to detecting social information (difference between correlations, $t = 2.47$, $p = 0.007$; Fig. 7*c*). Percentage of “Unsure” responses did not correlate with externalizing symptoms for either Random ($r_s = 0.048$, $p = 0.17$) or Mental ($r_s = 0.01$, $p = 0.75$) animations. Together, these analyses support a link

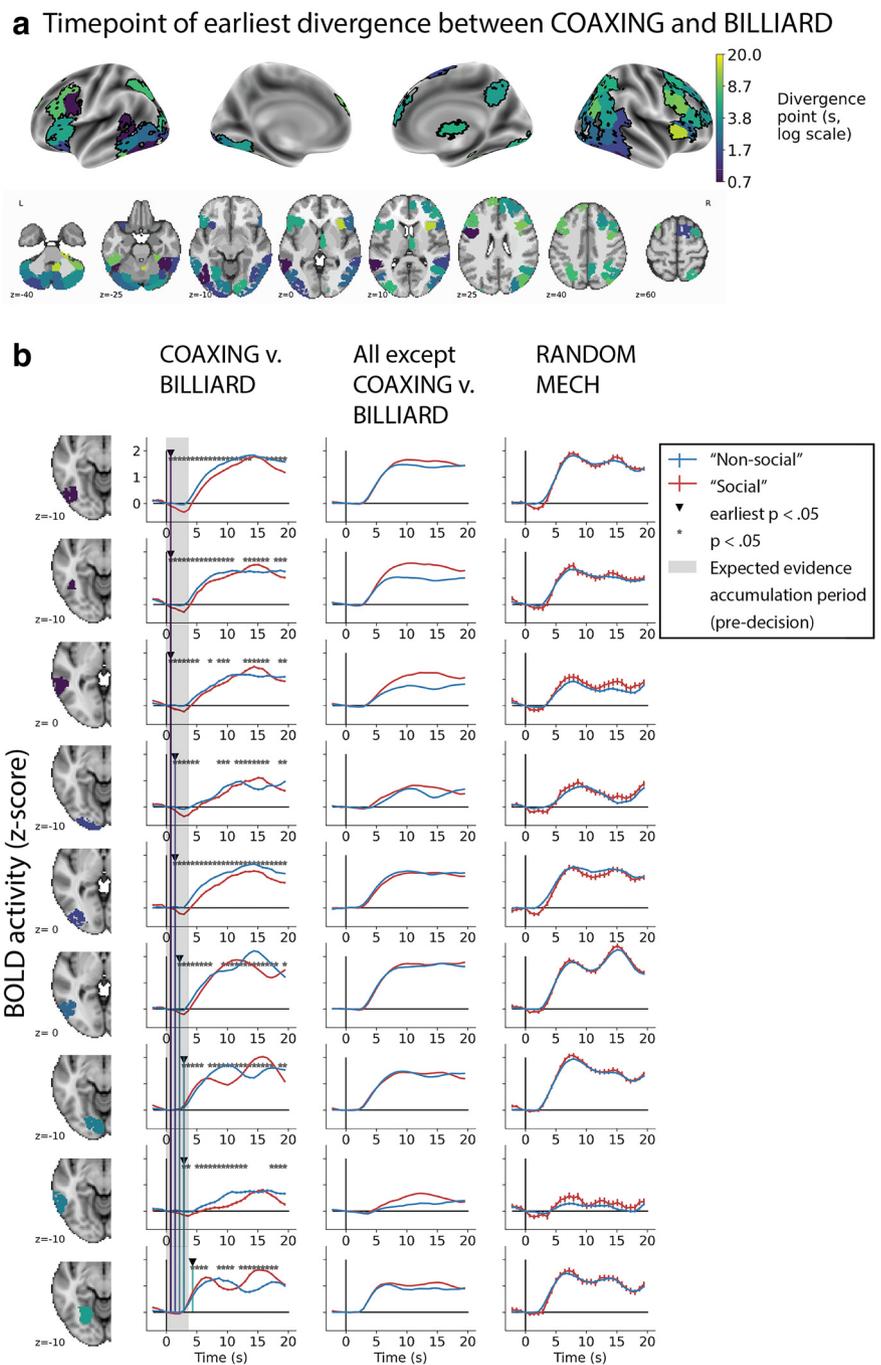


Figure 6. Time course analysis showing when and where differences between “Social” and “Non-social” percepts emerge. **a**, Brain map of the earliest time point at which brain activity diverges between “Social” and “Non-social” responses for the COAXING and BILLIARD animations, respectively (within-participant analysis). Analysis was limited to the robust social perception regions (compare Fig. 3, black contours), and BOLD signal time courses were residualized with respect to the visual features of brightness and optic flow to minimize the effects of any differences in low-level sensory information between the two animations. Colors show how early (purple-blue) or late (yellow-green) activity diverged. **b**, BOLD signal time courses in the left posterior regions illustrating how “Social” and “Non-social” activity diverge in the predecisional period for COAXING and BILLIARD. Regions are sorted in rows by the earliest divergence TR and then from posterior to anterior. Left, Time courses for the two animations matched for approximate decision time, COAXING (“Social”) and BILLIARD (“Non-social”). Middle and right, Time courses from the same regions from two supporting analyses, across all animations except COAXING–BILLIARD (“Social” vs “Non-social” response trials; middle); and for the most ambiguous animation, RANDOM MECH (“Social” vs “Non-social” responders), right.

between internalizing symptoms and a greater tendency to perceive information as social, perhaps driven by a homeostatic drive to seek social connections.

To understand whether overall neural activity while watching animations and scanning for social information also covaried

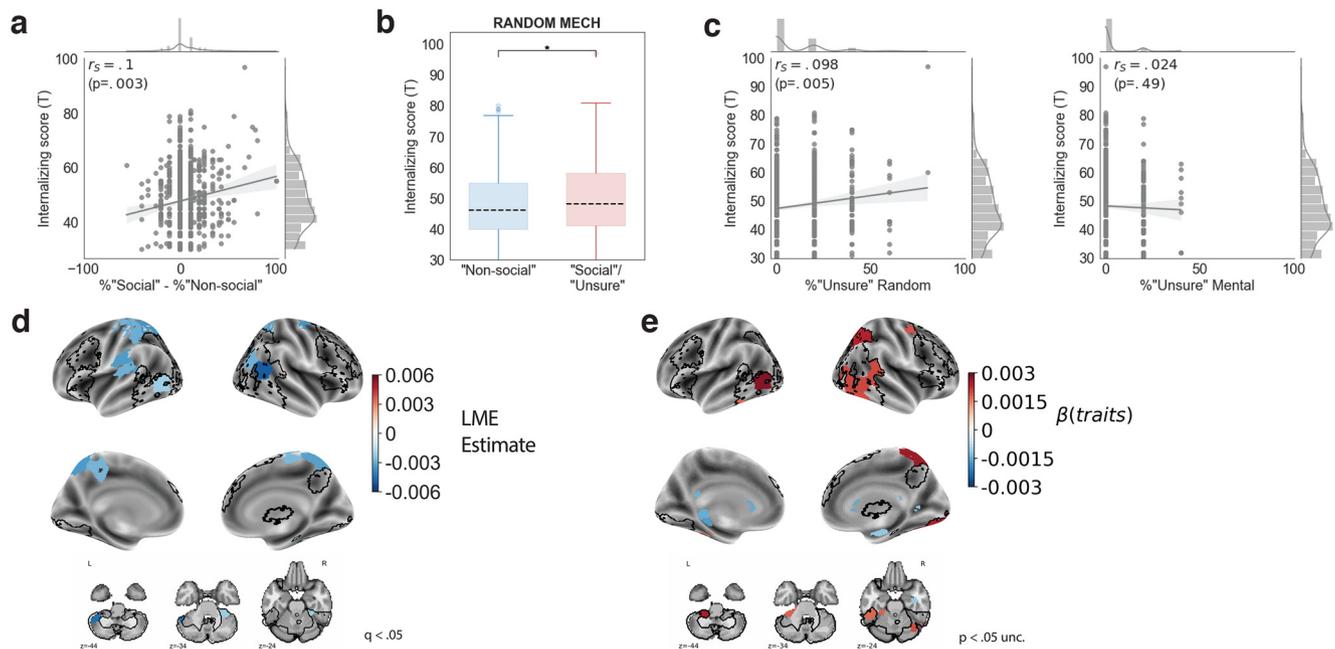


Figure 7. Relationship between internalizing trait scores, behavior, and brain activity. **a**, Response bias (percentage difference between “Social” and “Non-social” responses per participant) correlates positively with internalizing symptom score (Spearman’s correlation coefficient, $r_s = 0.1$, $p = 0.003$). **b**, Internalizing scores among individuals who reported some degree of socialness (“Social” or “Unsure” responses) to the most ambiguous animation, RANDOM MECH, were higher than those for individuals who reported this animation “Non-social”; $*p < 0.05$. **c**, Internalizing score correlates positively with the percentage of “Unsure” responses per participant for the generally non-social animations (Random, left, $r_s = 0.098$, $p = 0.005$) but not for the generally social animations (Mental, right, $r_s = -0.024$, $p = 0.49$). These correlation magnitudes were significantly different ($t = 2.47$, $p = 0.007$). **d**, Relationship between neural responsiveness during the task and internalizing scores. Colored parcels showed a significant positive (red) or negative (blue) relationship (FDR $q < 0.05$) with internalizing score. The robust social perception regions from the GLM analysis (compare Fig. 3) are shown in black. All regions show a negative relationship between activation magnitude and internalizing symptoms, and there is only a partial overlap with the robust social perception regions. **e**, Parcels in which the difference between activity to “Social” and “Non-social” percepts (compare Fig. 3c) may be modulated by internalizing symptoms ($p < 0.05$, uncorrected). Red indicates that at higher internalizing symptoms, “Social” trials show a relatively higher response than “Non-social”, and blue indicates that at higher internalizing symptoms, “Social” trials show a relatively lower response than “Non-social” (blue).

with internalizing symptoms, we related trialwise brain activity estimates to internalizing symptom scores in an LMEM (fixed effect, internalizing score; random effect, animation). In a whole-brain analysis, 18 parcels showed a significant relationship (FDR $q < 0.05$, Fig. 7d) between internalizing score and neural responsiveness. In all of these, the LME estimates were negative, that is, as internalizing scores increased, brain activity decreased, although all 18 parcels showed above-baseline activity as evidenced by the positive regression coefficients ($\beta > 0$ for all parcels). Thus, although individuals with higher internalizing scores showed positive activity in these regions when scanning animations for social information, the magnitude of this activity was lower than in individuals with lower internalizing scores. These relationships were seen in the right angular gyrus, the bilateral superior parietal lobule, left supramarginal gyrus, regions along the dorsal midline, and anterior parts of the cerebellum (Fig. 7d, blue).

Interestingly, the lateral occipital parcels from the set of robust social perception regions (Fig. 7d, black contours) were not as prominent here, showing only a partial overlap (five parcels) with the parcels showing trait effects. In the overlapping parcels, which comprised bilateral occipitotemporal parcels and the cerebellum, individuals high on internalizing traits showed overall less reactivity in many brain regions while scanning the environment for social interactions. To reconcile this decrease in neural reactivity (Fig. 7d) with the observed increase in behavioral sensitivity (Fig. 7a–c), one interpretation is that these individuals have

a lower threshold for the amount of neural activity required to declare something “Social”. Yet another interpretation, based on the decrease in neural activity with internalizing symptoms in all trait-sensitive parcels and the observation that 72% of the trait-sensitive regions (13 parcels) lie outside the previously identified robust social perception regions, is that this reflects a general decrease in neural responsiveness with more internalizing symptoms.

Finally, we probed whether the difference in neural reactivity to information ultimately perceived as “Social” showed any trait dependence. No parcel showed a significant relationship between internalizing scores and the subject-level “Social” – “Non-social” β estimates at the corrected threshold across all 268 parcels (FDR $q < 0.05$). At the uncorrected threshold ($p < 0.05$), however, we found several parcels (most notably in the right occipitotemporal parcels; Fig. 7e) showing a positive relationship in that individuals with higher internalizing symptoms showed a relatively higher responsiveness in this parcel to information eventually declared “Social” ($\beta_{\text{traits}} > 0$). Together with the overall effect of traits on brain activity, these results may suggest that individuals exhibiting higher internalizing symptoms show lower brain activity when scanning for social content but that the magnitude of this dip may be lower when viewing social content. However, this potential relationship should be further tested in datasets with more—and ideally more ambiguous—stimuli to allow for more variation in both behavioral and neural responses.

Discussion

In this study, we investigated behavioral and neural signatures of social signal detection using a large dataset of neurotypical young adults. Behavioral responses showed a subtle but consistent bias toward perceiving information as social (as opposed to non-social), which manifested as a higher number of “Social” responses and a hesitation to report information as “Non-social”. We then used observers’ own responses to label fMRI data and found that widespread patterns of brain activity differentiate conscious social percepts, even when controlling for visual input (RANDOM MECH) and decision time (COAXING–BILLIARD). Overall, observer responses explained more variance in activity than experimenter-assigned labels. Several regions also showed parametric responses to degrees of perceived socialness (“Social” greater than “Unsure” greater than “Non-social” responses). Further, brain activity for information ultimately deemed “Social” diverged from “Non-social” early both in time and in the cortical hierarchy. Finally, we found that a trait-level measure of internalizing symptoms (e.g., loneliness, anxiety) could explain some of the variability in percepts and brain activity.

Humans are an obligate social species predisposed toward social interactions (Epley et al., 2007; Rutherford and Kuhlmeier, 2013), and socially relevant content is processed more efficiently (Rothkirch et al., 2015; Papeo et al., 2017). The response bias toward “Social” in the current study, and its covariation with internalizing symptoms, supports the idea of a homeostatic drive to seek social connection (Tomova et al., 2020). This is in line with previous studies reporting that lonely people tend to form illusory social connections (Epley et al., 2008), overattribute animacy to faces (Powers et al., 2014), and have greater attention and memory for social cues (Gardner et al., 2005). It is also possible that the response bias observed here could have been partially because of the task structure, including contextual effects, instructions, or the very presence of multiple agents, which can induce expectations for social content (Piejka et al., 2022). However, it is unlikely that these factors fully account for our results. Despite their fixed order, trials were pseudorandomized with respect to experimenter-assigned labels, meaning that trial order should not have induced a systematic bias in responses. Furthermore, *all* animations (i.e., not just Mental ones) contained at least two agents, and the Mental and Random animations were largely matched in terms of the number and type of agents (Table 1).

Animated shapes have been used extensively in fMRI studies to characterize brain activity involved in social perception. Past work has converged on a canonical set of regions including bilateral pSTS, lateral occipital cortex, angular gyrus, superior parietal lobule, and medial prefrontal cortex (Castelli et al., 2000; Tavares et al., 2008; Osaka et al., 2012). However, nearly all this work has used stimuli generated by experimenters to be seen as obviously social or obviously non-social, then characterized participants’ “accuracy” with respect to these labels. There are two major limitations to this approach; one, differences in neural responses to social content may be confounded by differences in low-level visual features (e.g., higher speed for animations labeled social), and two, the idea that experimenter labels represent the ground truth is likely unrealistic given that real-world social scenarios are frequently ambiguous, and interpretations vary across individuals. Here, we extended the social perception literature in an important way: we eschewed experimenter-assigned labels and characterized brain activity according to participants’ own reported percepts, which allowed us to identify regions that are sensitive

to the subjective (or conscious) perception of a social interaction over and above sensory inputs. Several of these regions also showed parametric responses to degrees of perceived socialness, suggesting an even tighter link between activity in these regions and conscious perceptual experiences.

Because the animations used here and in and past work are typically relatively long in duration (20–40 s), another open question is to what extent the observed brain activity reflects distinct cognitive processes over the course of each trial. For example, early in animations, participants are likely accumulating evidence in favor of each alternative (i.e., “Social” or “Non-social”) until a decision is reached. Following this, postdecisional processes likely come into play, which could include maintaining the decision in working memory and monitoring for any counterevidence. Animations that have been deemed “Social” may enjoy higher levels of attention and engagement through the remainder of the trial, which could partially explain current and past observations (Tavares et al., 2008) of stronger neural responses to social content. Although the heterogeneity in decision times across individuals and animations (compare Fig. 2c) makes it challenging to disambiguate these processes, by leveraging stimuli with comparable decision times, we showed that several occipitotemporal regions start responding differently to information ultimately perceived as “Social” even before participants had likely arrived at a decision, suggesting that this activity may reflect predecisional evidence accumulation. This is further supported by the results of our parametric analysis, as well as recent EEG work showing a temporal hierarchy in action perception from encoding visual to social features (Dima et al., 2022). Early differences also emerged in the pSTS, an area critical to the third visual stream hypothesis (Pitcher and Ungerleider, 2021), and lateral parts of the precentral gyrus and the supplementary motor area. Recent electrophysiological studies (Isik et al., 2020; Dima et al., 2022) have shown that neural responses to social interactions occur at 300 ms and that socioaffective features best predict neural responses at ~418 ms, that is, on a timescale congruent with top-down processing. Thus, early activity in frontal regions may reflect feedback mechanisms that direct attentional resources in sensory cortex to prioritize processing social information.

In our trait-dependent analyses, we found that brain activity during the task was lower for individuals with high internalizing scores in regions including parts of the default mode network (angular gyrus and precuneus) and some of our previously identified robust social perception regions (occipitotemporal, frontal, cerebellar). Although this may reflect a decrease in neural reactivity while scanning the environment specifically for social information, we cannot rule out that it may also reflect change in neural activity to any task in these individuals. The apparent discrepancy whereby individuals with higher internalizing scores show increased behavioral sensitivity but decreased neural activity to potentially social content could indicate that these individuals have a lower neural threshold for declaring something “Social”. However, we found only weak evidence for interactions between internalizing scores, neural responses, and reported percepts. Future work might return to this question using more ambiguous stimuli that evoke more variability in both neural and behavioral responses across people.

One limitation of this dataset is that the stimulus set consisted of only 10 animations that were not counterbalanced in order across participants nor controlled in terms of their visual features. These animations were also not optimal to study ambiguous perception because all did have dominant percepts, although

the large sample size still allowed us to leverage nondominant percepts to separate conscious social percepts from sensory input (i.e., RANDOM MECH analysis). Future studies should replicate and extend these results using stimuli that are better controlled and also more ambiguous (i.e., evoke more balanced responses).

Another limitation is that participants were limited to three discrete response options, when perceptual certainty may have varied even within each response type. Furthermore, even for the same animation, individuals who perceive a social interaction do not necessarily perceive the same type of social interaction, and different interpretations could have muddied group-level effects. Future experiments can overcome this limitation by using richer behavioral characterizations of percepts.

In summary, we describe behavioral and neural processes that underlie how people arrive at conscious percepts of social information. Together, our results compel a more nuanced view of social perception in which socialness is in the eye of the beholder.

References

- Abassi E, Papeo L (2020) The representation of two-body shapes in the human visual cortex. *J Neurosci* 40:852–863.
- Abassi E, Papeo L (2022) Behavioral and neural markers of visual configural processing in social scene perception. *Neuroimage* 260:119506.
- Abell F, Happé F, Frith U (2000) Do triangles play tricks? Attribution of mental states to animated shapes in normal and abnormal development. *Cogn Dev* 15:1–16.
- Achenbach TM, Ivanova MY, Rescorla LA (2017) Empirically based assessment and taxonomy of psychopathology for ages 1½–90+ years: developmental, multi-informant, and multicultural findings. *Compr Psychiatry* 79:4–18.
- Banks WP (1970) Signal detection theory and human memory. *Psychol Bull* 74:81–99.
- Barch DM, Burgess GC, Harms MP, Petersen SE, Schlaggar BL, Corbetta M, Glasser MF, Curtiss S, Dixit S, Feldt C, Nolan D, Bryant E, Hartley T, Footer O, Bjork JM, Poldrack R, Smith S, Johansen-Berg H, Snyder AZ, Van Essen DC (2013) Function in the human connectome: task-fMRI and individual differences in behavior. *Neuroimage* 80:169–189.
- Barrett HC, Todd PM, Miller GF, Blythe PW (2005) Accurate judgments of intention from motion cues alone: a cross-cultural study. *Evol Hum Behav* 26:313–331.
- Blakemore S-J, Boyer P, Pachot-Clouard M, Meltzoff A, Segebarth C, Decety J (2003) The detection of contingency and animacy from simple animations in the human brain. *Cereb Cortex* 13:837–844.
- Castelli F, Happé F, Frith U, Frith C (2000) Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns. *Neuroimage* 12:314–325.
- Castelli F, Frith C, Happé F, Frith U (2002) Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain* 125:1839–1849.
- Davis JW, Gao H (2004) An expressive three-mode principal components model for gender recognition. *J Vis* 4(5):2 362–377.
- Deen B, Koldewyn K, Kanwisher N, Saxe R (2015) Functional organization of social perception and cognition in the superior temporal sulcus. *Cereb Cortex* 25:4596–4609.
- Desai A, Foss-Feig JH, Naples AJ, Coffman M, Trevisan DA, McPartland JC (2019) Autistic and alexithymic traits modulate distinct aspects of face perception. *Brain Cogn* 137:103616.
- Dima DC, Tomita TM, Honey CJ, Isik L (2022) Social-affective features drive human representations of observed actions. *Elife* 11:e75027.
- Epley N, Waytz A, Cacioppo JT (2007) On seeing human: a three-factor theory of anthropomorphism. *Psychol Rev* 114:864–886.
- Epley N, Akalis S, Waytz A, Cacioppo JT (2008) Creating social connection through inferential reproduction: loneliness and perceived agency in gadgets, gods, and greyhounds. *Psychol Sci* 19:114–120.
- Friston KJ, Jezzard P, Turner R (1994) Analysis of functional MRI time-series. *Hum Brain Mapp* 1:153–171.
- Gardner WL, Pickett CL, Jefferis V, Knowles M (2005) On the outside looking in: loneliness and social monitoring. *Pers Soc Psychol Bull* 31:1549–1560.
- Glasser MF, Sotiropoulos SN, Wilson JA, Coalson TS, Fischl B, Andersson JL, Xu J, Jbabdi S, Webster M, Polimeni JR, Van Essen DC, Jenkinson M (2013) The minimal preprocessing pipelines for the Human Connectome Project. *NeuroImage* 80:105–124.
- Glover GH (1999) Deconvolution of impulse response in event-related BOLD fMRI. *Neuroimage* 9:416–429.
- Gottsdanker R (1982) Age and simple reaction time. *J Gerontol* 37:342–348.
- Hebart MN, Donner TH, Haynes J-D (2012) Human visual and parietal cortex encode visual choices independent of motor plans. *Neuroimage* 63:1393–1403.
- Heider F, Simmel M (1944) An experimental study of apparent behavior. *Am J Psychol* 57:243–259.
- Isik L, Koldewyn K, Beeler D, Kanwisher N (2017) Perceiving social interactions in the posterior superior temporal sulcus. *Proc Natl Acad Sci U S A* 114:E9145–E9152.
- Isik L, Mynick A, Pantazis D, Kanwisher N (2020) The speed of human social interaction perception. *Neuroimage* 215:116844.
- Johnson KL, Tassinari LG (2005) Perceiving sex directly and indirectly: meaning in motion and morphology. *Psychol Sci* 16:890–897.
- Jolly E (2018) Pymer4: connecting R and Python for linear mixed modeling. *JOSS* 3:862.
- Kana RK, Maximo JO, Williams DL, Keller TA, Schipul SE, Cherkassky VL, Minshew NJ, Just MA (2015) Aberrant functioning of the theory-of-mind network in children and adolescents with autism. *Mol Autism* 6:59.
- Kanai R, Bahrami B, Duchaine B, Janik A, Banissy MJ, Rees G (2012) Brain structure links loneliness to social perception. *Curr Biol* 22:1975–1979.
- Landsiedel J, Daughters K, Downing PE, Koldewyn K (2022) The role of motion in the neural representation of social interactions in the posterior temporal cortex. *Neuroimage* 262:119533.
- Lee SM, Gao T, McCarthy G (2014) Attributing intentions to random motion engages the posterior superior temporal sulcus. *Soc Cogn Affect Neurosci* 9:81–87.
- Lessard LM, Juvonen J (2018) Friendless adolescents: do perceptions of social threat account for their internalizing difficulties and continued friendlessness? *J Res Adolesc* 28:277–283.
- Li G, Chen Y, Wang W, Dhingra I, Zhornitsky S, Tang X, Li C-SR (2020) Sex differences in neural responses to the perception of social interactions. *Front Hum Neurosci* 14:565132.
- Lisoy RS, Biegler R, Haghish EF, Veckenstedt R, Moritz S, Pfuhl G (2022) Seeing minds—a signal detection study of agency attribution along the autism-psychosis continuum. *Cogn Neuropsychiatry* 27:356–372.
- McNamara Q, De La Vega A, Yarkoni T (2017) Developing a comprehensive framework for multimodal feature extraction. [arXiv:1702.06151](https://arxiv.org/abs/1702.06151).
- McNemar Q (1947) Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika* 12:153–157.
- Mohammadzadeh A, Tehrani-doost M, Banaraki AK (2012) Evaluation of ToM (intentionality) in primary school children using movement shape paradigm. *Procedia Soc Behav Sci* 32:69–73.
- Nguyen M, Vanderwal T, Hasson U (2019) Shared understanding of narratives is correlated with shared neural responses. *Neuroimage* 184:161–170.
- Osaka N, Ikeda T, Osaka M (2012) Effect of intentional bias on agency attribution of animated motion: an event-related fMRI study. *PLoS One* 7:e49053.
- Palan S, Schitter C (2018) Prolific.ac—A subject pool for online experiments. *J Behav Exp Financ* 17:22–27.
- Palmer CJ, Clifford CWG (2020) Face pareidolia recruits mechanisms for detecting human social attention. *Psychol Sci* 31:1001–1012.
- Papeo L (2020) Twos in human visual perception. *Cortex* 132:473–478.
- Papeo L, Stein T, Soto-Faraco S (2017) The two-body inversion effect. *Psychol Sci* 28:369–379.
- Petrini K, McAleer P, Neary C, Gillard J, Pollick FE (2014) Experience in judging intent to harm modulates parahippocampal activity: an fMRI study with experienced CCTV operators. *Cortex* 57:74–91.
- Piejka A, Piaskowska L, Okruszek Ł (2022) Two means together? Effects of response bias and sensitivity on communicative action detection. *J Nonverbal Behav* 46:281–298.
- Pitcher D, Ungerleider LG (2021) Evidence for a third visual pathway specialized for social perception. *Trends Cogn Sci* 25:100–110.

- Powers KE, Worsham AL, Freeman JB, Wheatley T, Heatherton TF (2014) Social connection modulates perceptions of animacy. *Psychol Sci* 25:1943–1948.
- Rasmussen CE, Jiang YV (2019) Judging social interaction in the Heider and Simmel movie. *Q J Exp Psychol (Hove)* 72:2350–2361.
- Rothkirch M, Madipakkam AR, Rehn E, Sterzer P (2015) Making eye contact without awareness. *Cognition* 143:108–114.
- Rutherford MD, Kuhlmeier VA (2013) Social perception: detection and interpretation of animacy, agency, and intention. Cambridge, MA: MIT.
- Sacco DF, Merold SJ, Lui JHL, Lustgraaf CJN, Barry CT (2016) Social and emotional intelligence moderate the relationship between psychopathy traits and social perception. *Pers Individ Dif* 95:95–104.
- Schafroth JL, Basile BM, Martin A, Murray EA (2021) No evidence that monkeys attribute mental states to animated shapes in the Heider–Simmel videos. *Sci Rep* 11:3050.
- Scholl BJ, Tremoulet PD (2000) Perceptual causality and animacy. *Trends Cogn Sci* 4:299–309.
- Shen X, Tokoglu F, Papademetris X, Constable RT (2013) Groupwise whole-brain parcellation from resting-state fMRI data for network node identification. *Neuroimage* 82:403–415.
- Stanislaw H, Todorov N (1999) Calculation of signal detection theory measures. *Behav Res Methods Instrum Comput* 31:137–149.
- Tavares P, Lawrence AD, Barnard PJ (2008) Paying attention to social meaning: an fMRI study. *Cereb Cortex* 18:1876–1885.
- Tomova L, Wang KL, Thompson T, Matthews GA, Takahashi A, Tye KM, Saxe R (2020) Acute social isolation evokes midbrain craving responses similar to hunger. *Nat Neurosci* 23:1597–1605.
- Tremoulet PD, Feldman J (2000) Perception of animacy from the motion of a single object. *Perception* 29:943–951.
- Van Essen DC, Smith SM, Barch DM, Behrens TEJ, Yacoub E, Ugurbil K (2013) The WU-Minn Human Connectome Project: an overview. *Neuroimage* 80:62–79.
- Wagenmakers E-J, Farrell S (2004) AIC model selection using Akaike weights. *Psychon Bull Rev* 11:192–196.
- Walbrin J, Koldewyn K (2019) Dyadic interaction processing in the posterior temporal cortex. *Neuroimage* 198:296–302.
- Walbrin J, Downing P, Koldewyn K (2018) Neural responses to visually observed social interactions. *Neuropsychologia* 112:31–39.
- Wheatley T, Milleville SC, Martin A (2007) Understanding animate agents: distinct roles for the social network and mirror system. *Psychol Sci* 18:469–474.
- Williams E, Chakrabarti B (2022) The integration of head and body cues during the perception of social interactions. *PsyArXiv*. <https://doi.org/10.31234/osf.io/qc3vz>.
- Wood N, Cowan N (1995) The cocktail party phenomenon revisited: how frequent are attention shifts to one's name in an irrelevant auditory channel? *J Exp Psychol Learn Mem Cogn* 21:255–260.