Predictive Action Perception from Explicit Intention Information in Autism

Matthew Hudson[1], Toby Nicholson, Anna Kharko, Rebecca McKenzie, Patric Bach[2]

Plymouth University, UK.

Running head: PREDICTIVE ACTION PERCEPTION IN AUTISM

[1]Corresponding author now affiliated with: School of Business, National College of Ireland, Mayor Street, IFSC, Dublin 1, Ireland. Email: matthew.hudson@ncirl.ie

[2]Now affiliated with: University of Aberdeen, William Guild Building, Kings College, Old Aberdeen, Aberdeen, United Kingdom.

Abstract

Social difficulties in Autism Spectrum Disorder (ASD) may originate from a reduced top-down modulation of sensory information that prevents the spontaneous attribution of intentions to observed behaviour. However, although autistic people are able to explicitly reason about others' mental states, the effect of abstract intention information on perceptual processes has remained untested. ASD participants (n = 23) and a neurotypical (NT) control group (n = 23) observed a hand either reaching for an object or withdrawing from it. Prior to action onset, the participant either instructed the actor to "Take it" or "Leave it", or heard the actor state "I'll take it" or "I'll leave it", which provided an explicit intention that was equally likely to be congruent or incongruent with the subsequent action. The hand disappeared before completion of the action and participants reported the last seen position of the tip of the index finger by touching the screen. NT participants exhibited a predictive bias in response to action direction (reaches perceived nearer the object, withdrawals perceived farther away), and in response to prior knowledge of the actor's intentions (nearer the object after "Take it", farther away after "Leave it"). However, ASD participants exhibited a predictive perceptual bias only in response to the explicit intentions, but not in response to the motion of the action itself. Perception in ASD is not immune from top-down modulation. However, the information must be explicitly presented independently from the stimulus itself, and not inferred from cues inherent in the stimulus.

Keywords: Autism Spectrum Disorder, Action Prediction, Predictive Coding, implicit/explicit metalizing, representational momentum

Predictive Action Perception from Explicit Intention Information in Autism

Difficulties in social communication and interaction, alongside restricted and repetitive behaviours and interests, form a core symptom cluster of autism spectrum disorder (ASD), affecting the initiation and reciprocation of non-verbal, emotional, and communicative interactions (American Psychiatric Association, 2013). There is an emerging consensus that these social difficulties may originate, in part, from a general overreliance on bottom-up sensory input, at the expense of top-down information (e.g., Palmer, Seth, & Hohwy, 2015; Van de Cruys, et al., 2014). Reduced top-down influence would particularly affect social interactions, as the hidden goals and intentions that drive others' observed behaviour need to be constantly disambiguated (e.g. Bach, Nicholson, & Hudson, 2014; Barresi & Moore, 1996; Hohwy & Palmer, 2014). In typically developing individuals, such top-down influences allow others' behavior to be seen not as simply the movement of the limbs, but in terms of the mental states that cause it, as if others' intentions, beliefs, and emotions were "drawn" onto the behaviour (for reviews, see Bach & Schenke, 2017; Gallagher, 2008; Teufel, Fletcher, & Davis, 2010). This direct perception of the goals and intentions that imbue observed behavior with meaning may be less evident in ASD, so that behaviour is perceived only in terms of what is explicitly visually available.

So far, research has however primarily tested situations in which the relevant high-level information was merely *implied* by the observed behaviours, and not explicitly given, and therefore needed to be spontaneously inferred from the action kinematics. This limitation is crucial because there is emerging evidence that while autistic people can reason about others' mental states when explicitly asked, they tend to not do so spontaneously (e.g., Apperly & Butterfill, 2009; Frith & Frith, 2008; Schaller & Rauh, 2017; Senju, 2010; Tager-Flusberg, 2001; Zalla, Labruyere, Clement, & Georgieff, 2010). Despite being able to differentiate actions based on explicitly available visual information, for example recognizing different actions or discriminating them from robotic or scrambled motion (e.g., Cusack, Williams, & Neri, 2015), autistic people find it particularly difficult when action perception requires *spontaneous inferences* of higher-level information, such as the goals implied by an action (Ganglmayer, Schuwerk, Sodian & Paulus, 2019; Hudson, Burnett, & Jellema, 2012), an actor's false beliefs (Schuwerk, Jarvers, Vuori, & Sodian, 2016; Senju et al., 2010), the emotions implied by biological motion (Centelles, Assaiante, Etchegoyhen, Bouvard, & Schmitz, 2013), or the intention behind communicative gestures (Von der Luhe et al., 2016).

To convincingly test whether ASD involves a generalized reduction in top-down processing, it needs to be established that these characteristics persist even if the relevant higher-level information is explicitly given, in a form that makes spontaneous inferences unnecessary. This study provides this crucial test. It utilizes a recent paradigm (Hudson, Nicholson, Ellis, & Bach, 2016; Hudson, Nicholson, Simpson, Ellis, & Bach, 2016) that reliably reveals how social perception in neurotypical individuals is biased by the integration of both explicit top-down intention information, and intention information merely implied by the actor's behaviour. In each trial, participants view, on a monitor, an actor's hand near an object and are given explicit information about the actor's goal, either by hearing the actors verbal statements (e.g., "I'll take it!", "I'll leave it!"), or by themselves instructing the actor ("Take it!", "Leave it!"). They then briefly view the actor either reach for or withdraw from the object, equally likely following the expectation or acting against it. The hand disappears mid-trajectory, and participants simply report, either on a touch screen, or using a probe comparison task, the last seen position of the hand. Across several studies, these perceptual judgments were reliably distorted by the goal implied by the action kinematics, so that a reach was perceived nearer the object than it really was and a withdrawal further away (the representational momentum effect: Freyd & Finke, 1984). Importantly, perceptual judgments were also biased by the actor's explicit intention statements. Participants reported hands further towards an object if the intention was to “take” the object and further away if to “leave” it. Crucially, the effect increased the more reliably the intention predicted the action (Hudson, Bach & Nicholson, 2018), consistent with an optimal integration of prediction information into one’s perceptual estimates (Ernst & Banks, 2002).

These studies show that, in neurotypical participants, action perception is enriched by top-down goal expectations derived from both explicit intention information and action kinematics (Bach et al., 2014; Hudson, McDonough, Edwards, & Bach, 2018; McDonough, Costantini, Hudson, & Bach, 2020; McDonough, Hudson, & Bach, 2019). Here, we used this task with autistic individuals, and a matched group of neurotypical individuals, who completed both the spoken instruction and heard intention variants of the experiment. This study directly tests whether action observation in ASD is generally immune to top-down information, or whether it is evident if the goals are provided explicitly through the intention statements. Moreover, comparing the spoken and heard version of the task will reveal whether those with ASD generally over-weigh self- compared to other-generated predictors, as some proposals suggest (Lawson, Rees & Friston, 2014).

Method

Participants

Neurotypical participants (n=28) were recruited from Plymouth University and wider community. ASD participants (n=28) were recruited from local community services in the region via intermediaries within and outside of Plymouth University. All participants had normal/corrected vision and gave written informed consent. All ASD participants had received a clinical diagnosis of Autism Spectrum Disorder, meeting either DSM-IV/V (APA, 1994; 2013) or ICD-10 (WHO, 1993) criteria. Sample size was constrained by participant availability but matched previous studies. After exclusions (see Results), our sample size of

23 in each group provides 80% power (G*Power: Faul, Erdfelder, Lang, & Buchner, 2007) to detect across-groups within-subject effects sizes of .42, within-groups effect sizes of .61, and between-subject effect sizes of 0.84. The smallest effect sizes of interest detectable with this paradigm (SESOI, Lakens, Scheel, & Isager, 2018) are .30, .43 and .59, respectively. The experiment was approved by Plymouth University ethics board.

Participants completed the Autistic Spectrum Quotient (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001), which is a 50-item self-administered questionnaire that measures the degree of autistic traits in clinical and neurotypical populations along five dimensions (social skills, attention switching, imagination, attention to detail, and communication). Initial feedback from potential participants and their support worker suggested that the four response options (definitely/slightly agree, definitely/slightly disagree) be condensed to two (agree, disagree). The binary scoring of 1 or 0 for each item of the AQ permitted this with no loss of variance in AQ scores. Participants also completed the Vocabulary and Matrix Reasoning subtests (FSIQ-2) of the WASI IQ test (Wechsler, 1999), although we were unable to collect data from two NT participants. After participant exclusions (see below), the two groups did not differ in age, gender, Full-2 IQ, Verbal *t*-score or Perception *t*-score, whilst AQ scores were higher for the ASD group (see Table 1).

Table 1. Participant information after exclusions. Sample size and gender composition, with mean and SD of age, WASI score, and Autistic Spectrum Quotient Score for each group.

	N (females)	**Age**	**Left Handed**	**IQ**	**Verbal**	**Perception**	**AQ**
NT	23 (7)	40.7 (8.6)	8%	112.5 (14.0)	55.6 (11.3)	58.2 (7.1)	17.1 (7.8)
ASD	23 (5)	40.1 (12.5)	11%	106.5 (19.1)	50.8 (13.8)	55.7 (12.0)	30.2 (9.4)
P	.502	.848	.773	.246	.219	.410	<.001

Apparatus & Stimuli

The experiment was administered using Presentation (NeuroBS) on a Hewlett Packard s230tm EliteDisplay Touchscreen (1920x1080, 60 Hz). A Logitech PC120 combined microphone and earphone headset delivered auditory stimuli and registered verbal responses.

The stimulus set consisted of 8 action sequences, each consisting of 26 frames (960 X 540 px). They were filmed with a Canon Legria HFS200 at 30fps and separated into individual frames with MovieDek. Each sequence showed a hand reaching from the right to the left towards a goal object, depicting the initial start of the action but stopping before object contact. Sequences were digitally manipulated using Corel Paintshop Pro X6, so that all background details were replaced with a uniform black background. Four sequences showed natural reaches towards objects that were safe to touch (water bottle, water glass, wineglass, handle of a knife). In the other four sequences, the safe objects were digitally replaced by a painful object of similar size (cactus, broken water glass, broken wineglass, knife blade), ensuring identical reach kinematics in both conditions.

Each sequence started with the hand at a neutral middle point (randomly chosen between frames 11-16) and progressed forwards or backwards through the sequence in two-frame steps for 3, 4, or 5 frames (80ms each) to depict reaches for or withdrawals from the object. The response stimuli were created by erasing the hand from a single frame of each of the sequences to depict just the object. When displayed directly after the action, it created the impression that the hand simply disappeared.

Audio stimuli of an actor saying “I’ll take it” and “I’ll leave it” (1000ms duration) were recorded with a M-Audio Microtrack 2 Digital Voice Recorder, and were presented binaurally but biased 50% to the right ear to match the implied location of the actor.

Procedure

Participants first completed two training sessions, which served to progressively familiarise them with the task requirements (see Supplementary Materials for analysis). In the first training session (42 trials), participants reported, via touchscreen, the position of a briefly presented static hand. In the second training session (48 trials), participants judged the final position of a moving hand after it had disappeared. As in the main experiment, participants were instructed to touch the tip of the index finger.

Participants then completed two blocks of the experiment. Each trial (Figure 1) started with a written instruction to hold the spacebar with the right hand, to prevent tracking of the stimulus with the finger, while the other hand rested on the desk or lap. A fixation cross was then presented for 500 to 1000ms, followed by the first frame of the stimulus sequence (hand in neutral position). In separate blocks, participants either verbally instructed the actor (Spoken Instruction) or heard the actor state an intention (Heard Intention). In the Spoken Instruction block, participants said “Take it!” if the object was safe to grasp and “Leave it!” if the object was potentially painful. Responses were not recorded but correct performance was monitored throughout by the experimenter. In the Heard Intention block, participants heard the actor say “I’ll take it” if the object was safe, and “I’ll leave it” if the object was potentially painful. Block order was counterbalanced between participants. In both blocks, there was a 50% congruency between intention statement and action, and participants were explicitly told that the actor was just as likely to do what they had said/were instructed as do the opposite. The biases we measure therefore reflect people’s prior expectation that other’s statements truthfully signal their intention, rather than contingencies learned within the experiment.

In both conditions, the action sequence began 1200ms after the onset of the intention statement or the registration of the verbal response by Presentation’s sound threshold logic, showing the hand reaching for or withdrawing from the object, for 3, 4 or 5 frames. When the hand disappeared mid-action, participants released the spacebar and, with the same hand, touched the screen where the last seen position of the tip of the index finger had been. For each experimental session, participants completed 10 practice trials, followed by two blocks of 48 trials, providing all factorial combinations of object type (safe, painful), object identity (4), action direction (reach, withdrawal), and sequence length (3, 4, 5 frames).

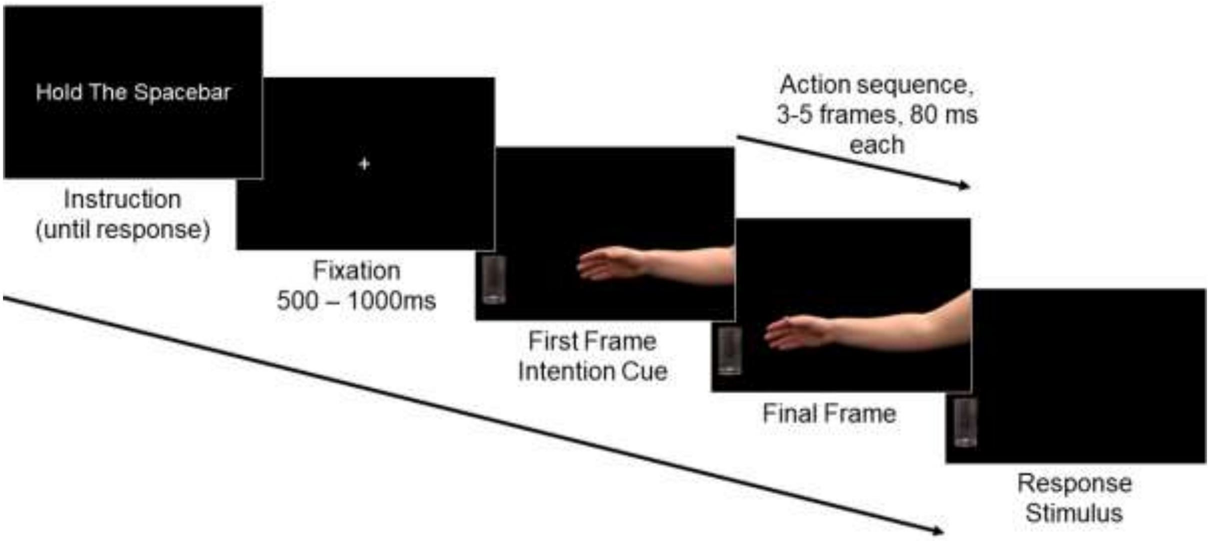

Figure 1. Trial sequence in the experimental blocks.

Trial and participant exclusion

Trials were excluded if the correct response procedure was not followed (participants released the spacebar prior to offset of the visual stimulus) and if the response initiation time (spacebar release after stimulus offset) or response execution time (spacebar release to touch screen response) was less than 200ms or more than 3SD above the group mean. Trial exclusions for the NT group (9%, SD=6.5%) and ASD group (14%, SD=15%) did not differ ($t(44)=1.61$, $p=.116$), leaving on average 174 trials per participant in the NT group (22 per condition) and 163 trials per participant in the ASD group (20 per condition). After exclusions, response initiations were slower to intention/action incongruent than congruent actions ($F(1,44)=7.52$, $p=.009$, $\eta p^2=.146$), and the ASD group was slower than the NT group to initiate responses ($F(1,44)=5.05$, $p=.03$, $\eta p^2=.103$) and execute responses ($F(1,44)=19.6$, $p<.001$, $\eta p^2=.308$), but this group effect did not interact with any of the conditions of interest (see Supplementary Materials). Participants were excluded (NT=4; ASD=3) if their responses showed no consistent relationship with the visual stimuli, as assessed by four a priori criteria (Hudson, Bach, et al., 2018): (1) if the selected screen position in each trial was not significantly correlated with the real screen position on either axis; (2) if the correlation coefficient between the real and selected screen positions was less than 3SD below the group median on either axis; (3) if the mean screen distance (pixels) between the real and selected screen positions was more than 3SD above the group mean; (4) if not enough valid trials (<50%) remained after trial exclusions (see below). One additional NT participant was excluded because they misunderstood the task, and two ASD participants were excluded for terminating the experiment before the second block.

Analysis

To assess the prediction that location judgments would be shifted in the direction of the actors' explicit goal (nearer the object after "take", farther after "leave") and in the direction of the action trajectory (nearer for reaches, farther for withdrawals), we measured the difference between the hands' location and where participants localized it on the touch screen. For each trial, the perceptual bias in pixels was calculated by subtracting the real disappearance point of the tip of the index finger from the participant's touch screen responses. Positive values reflect rightward (X axis) and upward (Y axis) biases from the real position, and negative values reflect leftward (X axis) and downward (Y axis) biases.

As in prior work (Hudson, Bach, et al., 2018; Hudson, McDonough, et al., 2018; McDonough et al., 2019), participants generally pressed further rightwards (34.0 px) and downwards (-105.3 px) of the real stimulus position, reflecting well-known shifts towards a stimulus' centre of gravity (Coren & Hoenig, 1972). The overall distance from the real position did not differ between groups ($t(44) = 1.51$, p=.140), but the ASD participants had a larger rightwards bias (ASD: 41.2px; NT: 26.9px, $t = 2.28$, $p = .027$), but a smaller downwards bias (ASD: -96.1px, $t = 2.57$, $p = .014$; NT: -114.5px). As these biases reflect general localization errors that are independent from our effects of interest, and are also present in the training sessions with dynamic and static stimuli (see Supplementary Material), the mean biases across groups were subtracted from each participant's mean response, therefore leaving all comparisons of interest unaffected.

These difference scores were analyzed separately for the X axis and Y axis with a mixed-effects ANOVA, with Intention Type (heard intention, spoken instruction), Action (reach, withdrawal), and Intention (take it, leave it) as within-subject factors, and Group (NT, ASD) as between-subjects factor. Main effects of Action will reveal the extent to which responses are biased away from the hands' real position towards the goals implied by the action kinematics (further leftwards towards the object for reaches than for withdrawals). Main effects of Intention will reveal the extent to which responses are biased away from the hands' real position towards the explicitly provided goal information (further leftwards for goals to "take" the object rather than "leave" it)[1]. Interactions with Group will reveal whether

[1] The main effect of intention is statistically equivalent to the congruency effect in which perceptual biases are larger for matching kinematics and prior intentions compared to mismatching ones. The difference between "Take" and "Leave" are free to differ (+/-) from zero independently of the action effect. The greater the difference between the intention conditions, the more maximal the perceptual shift leftwards is for reaches after "Take it" and the more maximal the perceptual shift rightwards is for withdrawals after "Leave it", relative to incongruent conditions (reaches after "Leave it", withdrawals after "Take it"). The intention effect can be reconceptualised as a non-directional shift (e.g., by sign reversing responses in the reach condition), so that they reflect perceptual shifts along the trajectory, irrespective of which direction that is. From that perspective, the disappearance point of reaches would generally be shifted forwards in the direction of motion, but less so when the intention was to "leave" rather than to "take" the object. Conversely, the disappearance point of withdrawals would be shifted forwards in the direction of motion, but less so when the intention was to "take" rather than to "leave" the object. Both analyses are statistically identical: they only differ in the sign with which deviations for reaches are coded (see also Hudson, Bach, et al., 2018).

ASD and NT groups differ in how much perceptual judgments are biased by kinematics (Action X Group) and explicit goal information (Intention X Group).

There were no between-group differences in variance on any of the effects of interest (Levene's max $F = 3.69$, $p = .061$). Bayes factors are reported for crucial two-sided tests with default priors using JASP (JASP Team, 2018), which provides a relative estimate of how much more likely the presence of an effect is given H1 than the null hypothesis. All other main effects or interactions will be evaluated against a Bonferroni-corrected alpha threshold for incidental findings due to hidden multiplicity in ANOVAs (Cramer, 2016).

Results

X axis

As predicted, and replicating our prior work, there was a main effect of Action, $F(1,44)=6.27$, $p=.016$, $\eta p^2=.13$, 95% CI [2.2, 23.7], BF10=2.18, showing that perceived disappearance points of reaches were shifted more leftward (towards the object) than those of withdrawals, and a main effect of Intention, $F(1,44)=53.4$, $p<.001$, $\eta p^2=.548$, 95% CI [6.9, 12.4], BF10=2.763e+6 (Figure 2), showing that perceptual judgements are biased more leftward after an explicit intention to "Take" the object than to "Leave" it, showing the predicted perceptual bias towards explicit action goals..

The main question was whether the perceptual biases induced by explicit intention information was smaller in the ASD group, as predicted from a generally reduced reliance of top-down information. However, there was strong evidence against an interaction of Intention and Group, $F(1,44)=1.78$, $p=.190$, $\eta_p^2=.04$, 95% CI [-8.9, 1.8], BF10=0.60. Separate analyses provided decisive evidence that intentions to "take" or "leave" the object biased perceptual judgments towards and away from this object in both groups, with similar effect sizes in the ASD, $t(22)=4.97$, $p<.001$, $d=1.03$, 95% CI [6.7, 16.2], BF10=451.7, and the NT group, *t*(22)=6.05, $p<.001$, *d* =1.26, 95% CI [5.2, 10.6], BF10=4809. Thus, explicit goal information biases social perception in both groups, refuting a generalized impairment in top-down processing.

Regarding motion prediction cues, the analysis revealed an interaction of Action and Group, $F(1,44)=4.30$, $p=.044$, $\eta_p^2=.09$, 95% CI [0.6, 42.3], BF10=1.61. Analysis of the NT group provided strong evidence that seeing reaches and withdrawals biased perceptual judgements towards or away from objects, $t(22)=4.43$, $p<.001$, $d=.92$, 95% CI [12.6, 34.7], BF10=141.5. However, for the ASD group, this analysis provided considerable evidence *against* such a bias, $t(22)=.25$, $p=.804$, $d=.05$, 95% CI [-20.6, 16.1], BF10=0.23. Thus, while perceptual judgments are biased by explicit intention information in ASD just as in NT participants, they are unaffected by goals implied by the action kinematics.

Finally, the analysis revealed the predicted interaction of Intention Type (spoken, heard) and Intention (take it, leave it), $F(1,44)=17.5$, $p<.001$, $\eta_p^2=.28$, 95% CI [3.1, 8.9], BF10=171.8, showing that spoken intentions induced larger perceptual shifts than heard intentions, but this did not differ by Group, $F(1,44)=1.2$, $p=.277$, $\eta_p^2=.03$, 95% CI [-2.7, 9.0], BF10=0.48. As we had no further predictions, all other effects are subject to alpha inflation in an ANOVA (Cramer et al., 2016, Bonferroni-adjusted $p<.003$). No further main effects or interactions passed the adjusted threshold.

Y axis

As we had no predictions for the Y-Axis, a Bonferroni-adjusted alpha of $p < .003$ was employed (Cramer et al., 2016). A main effect of Action, $F(1,44)=139.6$, $p<.001$, $\eta_p^2=.76$, 95% CI [31.8, 44.9], BF10=2.597e+12, showed that reaches were perceived higher than withdrawals, consistent with a bias in motion direction (i.e. upwards for reaches, downwards for withdrawals). There also was a main effect of Intention, $F(1,44)=28.6$, $p<.001$, $\eta_p^2=.394$, 95% CI [2.1, 4.5], BF10=7116. Hand disappearance points were perceived higher if the intention was to "Take" than to "Leave". There were no interactions between Intention and Group, $F(1,44)=.05$, $p=.825$, $\eta_p^2=.001$, 95% CI [-2.2, 2.8], BF10 = 0.30, or between Action and Group, $F(1,44)=1.4$, $p=.243$, $\eta_p^2=.03$, 95% CI [-5.4, 20.8], BF10=0.51.

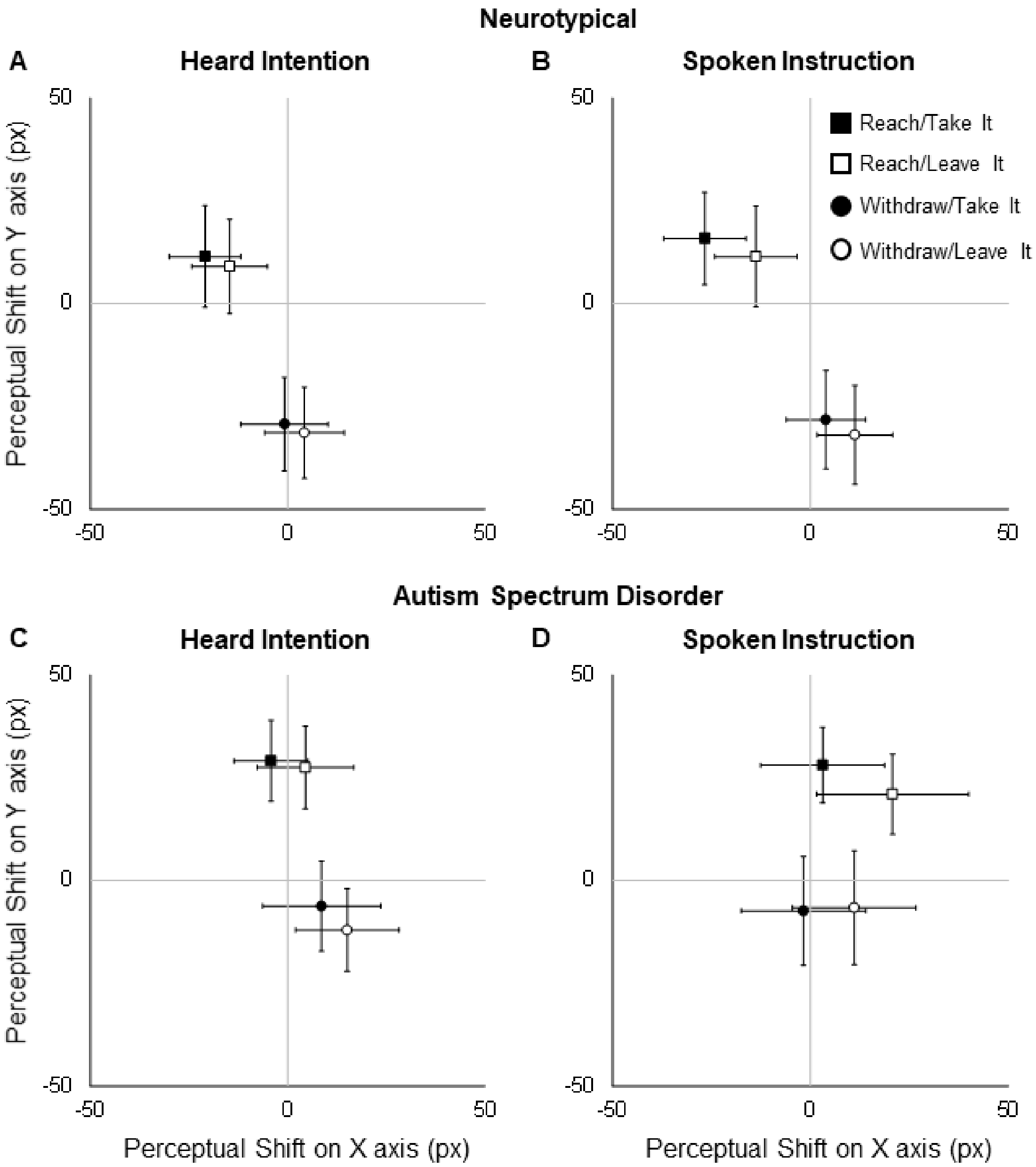

Figure 2. The perceptual bias (selected screen position – real screen position) on the X and Y axis for the NT (top row) and ASD (bottom row) groups, depending on whether explicit intentions were passively heard (left panels) or actively spoken (right panels). The intersection of the crosshairs at 0,0 represents the real disappearance point on any given trial, corrected for overall (across-group) biases in finger placement, with the object located to the left. In each panel, bias in the direction of the observed kinematics (main effect of Action) is

reflected in how much perceived disappearance points of reaches (squares) are shifted leftwards compared to those of withdrawals (circles). The bias in the direction of prior intentions (main effect of Intention) is reflected in how much explicit intentions to "take" the object (filled markers) shifts perceived disappearance points leftwards compared to intentions to "leave" the object (empty markers). Error bars represent 95% confidence intervals.

Relative weighting of motion and goal information

We next sought to establish whether the predictors of explicit intention and action kinematic information exert an equal influence on perceptual judgments, or whether one is weighted more than the other, and how this may differ between groups. For each participant, we derived statistically orthogonal indices of the effect of each predictor type by comparing the main effect of Action (withdrawals – reaches) and the main effect of Intention (Leave It – Take It). A mixed-effects ANOVA was conducted with Group (NT, ASD) as a between-subjects factor, and Intention Type (heard, spoken) and Predictor Type (motion, intention) as within-subjects factors (Figure 3A). The analysis revealed an interaction between Predictor and Group, $F(1,44)=4.71$, $p=.035$, $\eta_p^2=.097$, 95% CI [1.8, 48.2], BF10=1.88. The NT group showed a larger influence of motion ($m=23.7$px, $SD=25.6$) than intention information ($m=7.9$px, $SD=6.2$, $t(22) = 2.97$, $p = .007$, $d = 0.75$, 95% CI [4.7, 26.8], BF10=6.52). In contrast, the ASD group showed no bias, $t(22) =.902$, $p =.377$, $d = 0.34$, 95% CI [-30.4, 11.9], BF10=.315, although numerically there was a larger influence of intention information ($m=11.4$px, $SD=11.0$) than motion information ($m=2.2$px, $SD=42.5$).

There was a three-way interaction of Intention Type, Predictor, and Group, $F(1,44)=6.12$, $p=.017$, $\eta_p^2=.122$, 95% CI [5.7, 55.2], BF10=3.22. When analysed separately, the NT group provided considerable evidence for the absence of an interaction of Intention Type and Predictor, $F(1,22)=.78$, $p=.387$, $\eta_p^2=.034$, 95% CI [-12.7, 5.1], BF10=0.311. Perceptual judgments were weighted more in favour of motion information than intention information in both the Spoken and Heard conditions. This interaction was, however, present in the ASD group, $F(1,22)=5.34$, $p=.031$, $\eta_p^2=.95$, 95% CI [2.7, 50.5], BF10=1.96, who, unlike the NT group, weighted intention information more strongly than motion information in the Spoken condition but not the Heard condition.

Exploratory correlational analyses

We assessed the relationship between the weighting of motion and intention information across participants (Figure 3B). The NT group showed no correlation in the perceptual shift caused by motion and intention information, $r=.136$, $n=23$, $p=.536$, 95% CI [-.314, .585], BF10=0.31. However, in the ASD group, those participants who afforded less weight to motion information gave more weight to intention information, $r=-.503$, $n=23$, $p=.014$, 95% CI [-.895, -.111], BF10=4.33.

Finally, we tested whether the relative influence of motion and intention information was associated with individual differences in IQ or autistic traits, as measured by the AQ and its subscales. For each participant, the perceptual shift elicited by intention information was subtracted from that elicited by motion kinematics, indexing the size to which each participant weighted each predictive source. The NT group exhibited no relationship between this relative weighting and overall AQ score, nor any of its subscales, or the verbal or perception WASI sub-tests. The ASD group exhibited no relationship between autistic-like traits and the preference of motion or intention information except for a marginal (against Bonferroni-corrected $p = .008$) negative association with the social skills sub-scale, $r=-.538$, $p=.008$, 95% CI [-.155, -.920], BF10=7.0. The greater the deficit in social skills, the larger the influence of explicit intention information on the intention-action prediction effect. Moreover, there were significant relationships between the extent of motion information weighting and performance on the verbal, $r=.489$, $p=.018$, 95% CI [.094, .885], BF10=3.64, and perceptual subtests, $r=.523$, $p=.010$, 95% CI [.137, .910], BF10=5.70, indicating that the higher the IQ score, irrespective of domain, the greater the influence of motion information compared to intention information.

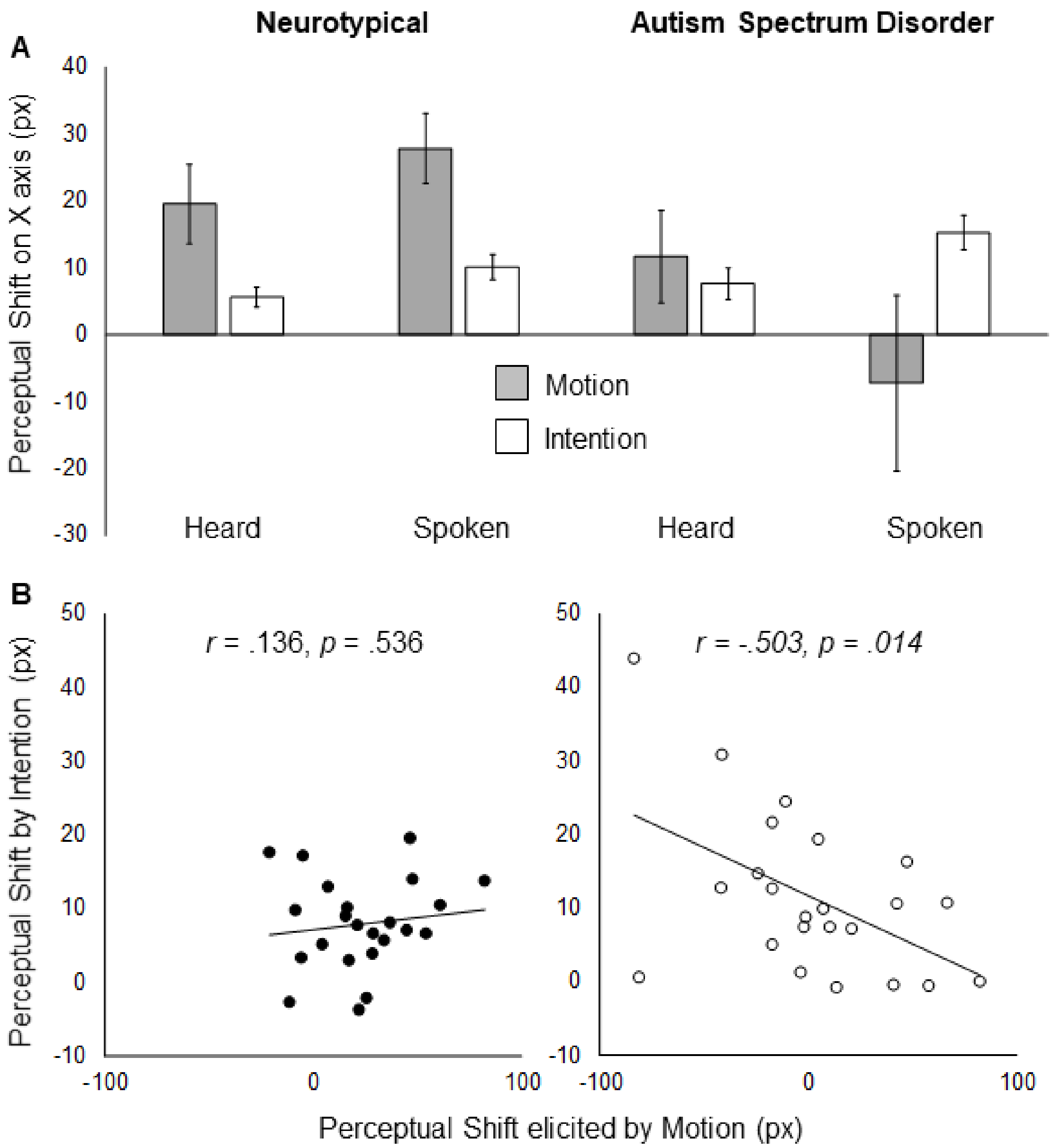

Figure 3. Panel A: The relative weighting of Motion information and Intention information on the perceptual shift for NT (left) and ASD (right) groups. Error bars represent standard error of the mean. Panel B: The relationship between the weighting of Motion and Intention information for each participant in the NT (left) and ASD (right) groups.

Discussion

Here, we tested whether social perception in autism spectrum conditions is impervious to top-down expectation (e.g., Palmer et al., 2015; Van de Cruys, et al., 2014). Replicating our prior work (Hudson, Bach, et al., 2018; Hudson, Nicholson, Ellis, et al., 2016; Hudson, Nicholson, Simpson, et al., 2016), in NT participants, perceptual judgments of others' actions were predictively biased by both the goals implied by the observed kinematics (a reach or withdrawal) and by the explicit intentions (to "take" or "leave" the object). The ASD participants differed markedly from this pattern, but in a manner inconsistent with a generally weaker top-down influence. The explicit intention statements to "take" or "leave" the object biased perceptual judgments as in NT participants, revealing that top-down influence of higher-level goal information onto action perception is fundamentally intact. However, ASD participants exhibited no perceptual shift in response to the goals implied by the observed movements. Thus, top-down expectations about another's goals shape social perception in ASD just as they do in NT individuals, but only if these expectations are explicitly provided and do not have to be spontaneously inferred from the observed kinematics itself.

These findings challenge the growing consensus that social perception in ASD is not influenced by top-down modulation (e.g., Palmer et al., 2015; Van de Cruys, et al., 2014). Instead, our findings imply a more differentiated account in which prediction channels for perceptually derived top-down information are down-weighted in ASD, while others remain intact, perhaps guided by the relative reliability of prediction sources (Ernst & Banks, 2002). Indeed, our analysis of the relative weighting of both types of intention cues found that neurotypical participants weighted prior motion information more strongly than intention information, as expected given that current motion is the more reliable predictor of future position. In contrast, for the ASD group, explicit intention information exerted a greater influence on the top-down predictive bias than action kinematics, specifically when these intentions were given to the actors by the participants (rather than passively heard). In addition, in ASD participants, but not in the NT group, the weighting of both predictors was negatively associated, with the influence of intention information increasing the more that of kinematic information decreased. This suggests a difficulty integrating the two sources of information, which is resolved by down-weighting kinematic relative to explicit intention information (Ernst & Banks, 2002; Hudson, Bach, et al., 2018; Zaki, 2013), although these results should be interpreted with caution due to the low sample sizes.

The question then is why autistic people would form predictions only when the predictive cue is made explicit. This may speak to the hierarchical organization of the social perception system. Higher-level predictions generalize across space and time and become translated into action expectations further down the hierarchy, possibly via goal-to-kinematic transformations in the motor system (Kilner, 2011), about what one will perceive, such as the assumption that others will behave in the most efficient way possible to achieve their goals (Marsh, Pearson, Ropar, & Hamilton, 2015). Action expectations can, however, also form locally, within the perceptual system for analyzing biological motion (Scholl & Gao, 2013; Hudson, McDonough, et al., 2018; McDonough et al., 2019). Our data suggest that it is in these low-level prediction processes that the differences in social perception may originate (Palmer, Lawson, & Hohwy, 2017). This could reflect either the use of kinematics to generate perceptual predictions about an action's future course, or the perception of a coherent motion percept from the sequential motion frames, which requires predictive processes (Kourtzi & Shiffrar, 1999; Yantis & Nakama, 1998), but which is compromised in ASD (David et al., 2010).

These findings of reduced reliance on motion prediction maps nicely onto wider issues in biological motion perception in ASD, whereby small reductions in biological motion sensitivity may give rise to up-stream difficulties in the interpretation of social information from motion cues, such as emotion (for meta-analyses see, Federici, Parma, Vicovaro, Radassao, Casartelli, & Ronconi, 2020; Todorova, McBean Hatton, & Pollick, 2019). Interestingly, motion perception sensitivity is related to the extent of the autistic condition (Kaiser & Shiffrar, 2009) and we similarly found that the underweighting of motion relative to intention information was negatively associated with IQ and social skills. While these findings need to be confirmed in more complex and realistic behaviors, they suggest that the shift from spontaneous inferences to explicit information could become more pronounced with a more heterogeneous sample that encompasses all levels of the autistic spectrum. Moreover, perceptual processes may be employed for prediction if provided with explicit information about the reliability of such sensory information (e.g., Hudson, Bach, et al., 2018).

The current findings have implications beyond the domain of social perception. Atypical social perception overlaps with that of non-social perception, reflecting a broader difference in integrating top-down input, resulting in both perceptual enhancements and impairments (Behrmann, Thomas, & Humphreys, 2006; Mottron, Dawson, Soulieres, Hubert, & Burack, 2006). An increased reliance on bottom-up sensory evidence at the expense of top-down information can explain many characteristics of perceptual anomalies in ASD, such as the preference for local details over global configuration (Happé & Frith, 2006), higher motion coherence thresholds (Milne, Swettenham, Hansen, Campbell, Jeffries, & Plaisted, 2002; Zaidel, Goin-Kochel, & Angelaki, 2015) or the reduced use of complex/second order motion cues (Bertone, Mottron, Jelenic, & Faubert, 2003; Kaiser & Shiffrar, 2009). We suggest that these findings may not reflect a general reduction in the use of high-level information to shape sensory input, but instead the spontaneous generation of the relevant inferences from the input.

Conclusion

Social perception in ASD is subject to top-down modulation but only if this information is explicitly given as abstract information that is independent of the stimulus itself. Such perceptual processes in the social domain may represent an especially notable example of a more general atypicality in ASD perception. This study opens up a new avenue for perceptual research in ASD. Perceptual processes in ASD may in fact be open to top-down influences if such information is explicitly provided.

Open Practices Statement

The experiment was not formally preregistered. De-identified raw and aggregate data for the experiments are posted at osf.io/k3m6n.

References

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). Arlington, VA: American Psychiatric Publishing.

Apperly, I.A. & Butterfill, S.A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review, 116*, 953-970. https://doi.org/10.1037/a0016923

Bach, P., Nicholson, T., & Hudson, M. (2014). The affordance-matching hypothesis: how objects guide action understanding and prediction. *Frontiers in Human Neuroscience, 8,* article 254. https://doi.org/10.3389/fnhum.2014.00254

Bach, P. & Schenke, K.C. (2017). Predictive social perception: Towards a unifying framework from action observation to person knowledge. *Social and Personality Psychology Compass, 11,* e12312. https://doi.org/10.1111/spc3.12312

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The Autism-Spectrum Quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders, 31,* 5-17. https://doi.org/10.1023/A:1005653411471

Barresi, J. & Moore, C. (1996). Intentional relations and social understanding. *Behavioral and Brain Sciences, 19,* 107-154. https://doi.org/10.1017/S0140525X00041790

Behrmann, M., Thomas, C., & Humphreys, K. (2006). Seeing it differently: Visual processing in autism. *Trends in Cognitive Sciences, 10,* 258-264. https://doi.org/10.1016/j.tics.2006.05.001

Bertone, A., Mottron, L., Jelenic, P. & Faubert, J. (2003). Motion perception in autism: A "complex" issue. *Journal of Cognitive Neuroscience, 15,* 218-225. https://doi.org/10.1162/089892903321208150

Centelles, L., Assaiante, C., Etchegoyhen, K., Bouvard, M., & Schmitz, C. (2013). From action to interaction: Exploring the contribution of body motion cues to social understanding in typical development and in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 43,* 1140–50. https://doi.org/10.1007/s10803-012-1655-0

Coren, S., & Hoenig, P. (1972). Effect of non-target stimuli upon length of voluntary saccades. Perception and Motor Skills, 34, 499–508. https://doi.org/10.2466/pms.1972.34.2.499

Cramer, A. O., van Ravenzwaaij, D., Matzke, D., Steingroever, H., Wetzels, R., Grasman, R. P., ... & Wagenmakers, E. J. (2016). Hidden multiplicity in exploratory multiway ANOVA: Prevalence and remedies. *Psychonomic Bulletin & Review*, *23,* 640-647. https://doi.org/10.3758/s13423-015-0913-5

Cusack, J.P., Williams, J.H.G., & Neri, P. (2015). Action perception is intact in autism spectrum disorder. *The Journal of Neuroscience, 35,* 1849-1857. https://doi.org/10.1523/JNEUROSCI.4133-13.2015.

David, N., Rose, M., Schneider, T.R., Vogeley, K., & Engel, A.K. (2010) Altered horizontal binding of single dots to coherent motion in autism. *Journal of Autism and Developmental Disorders, 40,* 1549-1551. https://doi.org/10.1007/s10803-010-1008-9

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

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191. https://doi.org/10.3758/BF03193146

Federici, A., Parma, V., Vicovaro, M., Radassao, L., Casartelli, L., & Ronconi, L. (2020)

Anomalous Perception of Biological Motion in Autism: A Conceptual Review and Meta-Analysis. *Scientific Reports, 10:*4576. https://doi.org/10.1038/s41598-020-61252-3

Freyd, J. J., & Finke, R. A. (1984). Representational momentum. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10,* 126-132. https://doi.org/10.1037/0278-7393.10.1.126

Frith, C.D. & Frith, U. (2008). Implicit and explicit processes in social cognition. *Neuron, 60,* 503-510. https://doi.org/10.1016/j.neuron.2008.10.032

Gallagher, S. (2008). Direct perception in the intersubjective context. *Consciousness and Cognition*, *17,* 535-543. https://doi.org/10.1016/j.concog.2008.03.003

Ganglmayer, K., Schuwerk, T., Sodian, B., & Paulus, M. (2019). Do children and adults with autism spectrum condition anticipate others' actions as goal-directed? A predictive coding perspective. *Journal of Autism and Developmental Disorders.* https://doi.org/10.1007/s10803-019-03964-8

Happé, F. & Frith, U. (2006). The weak central coherence account: Detail focussed cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 36,* 5-25. https://doi.org/10.1007/s10803-005-0039-0

Hohwy, J. & Palmer, C. (2014). Social cognition as causal inference: implications for common knowledge and autism. In Mattia Gallotti and John Michael (Eds.), *Social Ontology and Social Cognition*, Springer Series "Studies in the Philosophy of Sociality", Vol. 4.

Hudson, M., Bach, P., & Nicholson, T. (2018). You said you would! The predictability of other's behavior from their intentions determines predictive biases in action perception. *Journal of Experimental Psychology: Human Perception and Performance. 44,* 320-335. https://doi.org/10.1037/xhp0000451

Hudson, M., Burnett, H.G., & Jellema, T. (2012). Anticipation of action intentions in autism spectrum disorder. *Journal of Autism and Developmental Disorders, 42,* 1684-1693. https://doi.org/10.1007/s10803-011-1410-y.

Hudson, M, McDonough KL, Edwards R, Bach P. (2018). Perceptual teleology: expectations of action efficiency bias social perception. *Proceedings of the Royal Society: Series B, 285*: 20180638. http://dx.doi.org/10.1098/rspb.2018.0638

Hudson, M., Nicholson, T., Ellis, R. & Bach, P. (2016). I see what you say: Prior knowledge of other's goals automatically biases the perception of their actions. *Cognition, 146,* 245-250. https://doi.org/10.1016/j.cognition.2015.09.021

Hudson, M., Nicholson, T., Simpson, W., Ellis, R., & Bach, P. (2016). One step ahead: the perceived kinematics of others' actions are biased towards expected goals. *Journal of Experimental Psychology: General. 145,* 1-7. https://doi.org/10.1037/xge0000126

Kaiser, M.D. & Shiffrar, M. (2009). The visual perception of motion by observers with autism spectrum disorders: A review and synthesis. *Psychonomic Bulletin and Review, 16,* 761-777. https://doi.org/10.3758/PBR.16.5.761

Kilner, J. M. (2011). More than one pathway to action understanding. *Trends in Cognitive Sciences, 15,* 352–357. http://dx.doi.org/10.1016/j.tics.2011.06.005

Kourtzi, Z., & Shiffrar, M. (1999). Dynamic representations of human body movement. *Perception, 28,* 49–62. https://doi.org/10.1068/p2870

Lakens, D., Scheel, A.M., & Isager, P.M. (2018). Equivalence Testing for Psychological Research: A Tutorial. *Advances in Methods and Practices in Psychological Science, 1,* 259-269, https://doi.org/10.1177/2515245918770963

Lawson, R. P., Rees, G. & Friston, K.J. (2014). An aberrant precision account of autism. *Frontiers in Human Neuroscience, 8,* article 302. doi:10.3389/fnhum.2014.00302

Marsh, L.E., Pearson, A., Ropar, D. & Hamilton, A.F. de C. (2015). Predictive gaze during observation of irrational actions in adults with autism spectrum conditions. *Journal of Autism and Developmental Disorders, 45,* 245-261. https://doi.org/10.1007/s10803-014-2215-6

McDonough, K.L., Costantini, M., Hudson, M., & Bach, P. (2020). Affordance matching predictively shapes the perceptual representation of others' ongoing actions. Journal of Experimental Psychology: Human Perception and Performance, 46(8), 847-859. http://dx.doi.org/10.1037/xhp0000745

McDonough, K. L., Hudson, M., & Bach, P. (2019). Cues to intention bias action perception toward the most efficient trajectory. *Scientific Reports, 9*:6472, 1-10. https://doi.org/10.1038/s41598-019-42204-y

Milne, E., Swettenham, J., Hansen, P., Campbell, R., Jeffries, H. & Plaisted, K. (2002). High motion coherence thresholds in children with autism. *Journal of Child Psychology and Psychiatry,* 43, 255-263. https://doi.org/10.1111/1469-7610.00018

Mottron, L., Dawson, M., Soulieres, I., Hubert, B., & Burack, J. (2006). Enhanced perceptual functioning in autism: An update and eight principles of autism perception. *Journal of Autism and Developmental Disorders, 36,* 27-43. https://doi.org/10.1007/s10803-005-0040-7

Palmer, C.J., Lawson, R.P. & Hohwy, J. (2017). Bayesian approaches to autism: Towards volatility, Action, and Behavior. *Psychological Bulletin, 143,* 521-542. https://doi.org/10.1037/bul0000097.

Palmer, C.J., Seth, A.K. & Hohwy, J. (2015). The felt presence of other minds: Predictive processing, counterfactual predictions, and mentalising in autism. *Consciousness and Cognition, 26,* 376-389. https://doi.org/10.1016/j.concog.2015.04.007

Schaller, U.M. & Rauh, R. (2017). What difference does it make? Implicit, explicit and complex social cognition in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 47,* 961-979. https://doi.org/10.1007/s10803-016-3008-x

Scholl, B.J. & Gao, T. (2013). Perceiving animacy and intentionality: Visual processing or higher-level judgment? In Social perception: detection and interpretation of animacy, agency, and intention (eds MD Rutherford, VA Kuhlmeier), pp. 197–230. Cambridge, MA: MIT Press.

Schuwerk, T., Jarvers, I., Vuori, M. & Sodian, B. (2016). Implicit mentalizing persists beyond early childhood and is profoundly impaired in children with autism spectrum condition. *Frontiers in Psychology, 7,* article 1696. https://doi.org/10.3389/fpsyg.2016.01696

Senju, A., Southgate, V., Miura, Y., Matsui, T., Hasegawa, T., Tojo, Y., Osanai, H., & Csibra, G. (2010). Absence of spontaneous action anticipation by false belief attribution in children with autism spectrum disorder. *Developmental and Psychopathology, 22,* 353-360. https://doi.org/10.1017/S0954579410000106.

Tager-Flusberg, H. (2001). A re-examination of the Theory of Mind hypothesis of autism. In J. Burack, T. Charman, N. Yirmiya, & P. Zelazo (Eds.), The development of autism: Perspectives from theory and research, Mahwah, NJ: Lawrence Erlbaum Associates.

Teufel, C., Fletcher, P. C., & Davis, G. (2010). Seeing other minds: attributed mental states influence perception. *Trends in Cognitive Sciences, 14,* 376–82. https://doi.org/10.1016/j.tics.2010.05.005

Todorova, G.K., McBean Hatton, R.E., & Pollick, F.E. (2019). Biological motion perception in autism spectrum disorder: A meta-analysis. *Molecular Autism, 10:* 49. https://doi.org/10.1186/s13229-019-0299-8

Van de Cruys, S., Evers, K., Van der Hallen, R., Van Eylen, L., Boets, B., de-Wit, L. &

Wagemans, J. (2014). Precise minds in uncertain worlds: Predictive coding in autism. *Psychological Review, 121,* 649-675. https://doi.org/10.1037/a0037665.

von der Luhe, T., Manera, V., Barisic, I., Becchio, C., Vogeley, K. & Schilbach, L. (2016). Interpersonal predictive coding, not action perception, is impaired in autism. *Philosophical Transactions of the Royal Society: Series B, 371,* 20150373. https://doi.org/10.1098/rstb.2015.0373

Wechsler, D. (1999). Wechsler Abbreviated Scale of Intelligence. San Antonio, TX: The Psychological Corporation.

Yantis , S. & Nakama, T. (1998). Visual interactions in the path of apparent motion. *Nature Neuroscience, 6,* 508-512. https://doi.org/10.1038/2226

Zaidel, A., Goin-Kochel, R.P. & Angelaki, D.E. (2015). Self-motion perception in autism is compromised by visual noise but integrated optimally across multiple senses. *Proceedings of the National Academy of Sciences, 112,* 6461-6466. https://doi.org/10.1073/pnas.1506582112.

Zaki, J. (2013). Cue integration: A common framework for social cognition and physical perception. Perspectives of Psychological Science, 8, 296-312. https://doi.org/10.1177/1745691613475454

Zalla, T., Labruyere, N., Clement, A. & Georgieff, N. (2010). Predicting ensuing actions in children and adolescents with autism spectrum disorders. *Experimental Brain Research, 201,* 809-819. https://doi.org/10.1007/s00221-009-2096-7.

Author contributions

MH, TN, RM, PB devised the experiment. MH, TN, AK collected data. MH analysed the data, and wrote the MS with PB. All authors have approved the final manuscript.

Acknowledgments

Thanks to Charlotte Brind and Kiara Konstantinidou for WASI data collection, and to Barbara Wilson for participant liaison, recruitment, and organisation.

Funding

This work was supported by the Economic and Social Research Council (grant no. ES/J019178/1) awarded to P.B.