Spontaneous instrumental approach-avoidance learning in social contexts in autism

Morgan Beaurenaut[1*], Klara Kovarski[2,3], Constance Destais[4], Rocco Mennella[1,4†] and Julie Grèzes[4*†]

Abstract

Background Individuals with Autism Spectrum Condition (ASC) are characterized by atypicalities in social interactions, compared to Typically Developing individuals (TD). The social motivation theory posits that these difficulties stem from diminished anticipation, reception, and/or learning from social rewards. Although learning from socioemotional outcomes is core to the theory, studies to date have been sparse and inconsistent. This possibly arises from a combination of theoretical, methodological and sample-related issues. Here, we assessed participants' ability to develop a spontaneous preference for actions that lead to desirable socioemotional outcomes (approaching/avoiding of happy/angry individuals, respectively), in an ecologically valid social scenario. We expected that learning abilities would be impaired in ASC individuals, particularly in response to affiliative social feedback.

Method We ran an online social reinforcement learning task, on two large online cohorts with ($n = 274$) and without ($n = 290$) ASC, matched for gender, age and education. Participants had to indicate where they would sit in a waiting room. Each seat was associated with different probabilities of approaching/avoiding emotional individuals. Importantly, the task was implicit, as participants were not instructed to learn, and emotional expressions were never mentioned. We applied both categorical analyses contrasting the ASC and TD groups and dimensional factor analysis on affective questionnaires.

Results Contrary to our hypothesis, participants showed spontaneous learning from socioemotional outcomes, regardless of their diagnostic group. Yet, when accounting for dimensional variations in autistic traits, as well as depression and anxiety, two main findings emerged among females who failed to develop explicit learning strategies: (1) autism severity in ASC correlated with reduced learning to approach happy individuals; (2) anxiety-depression severity across both ASC and TD participants correlated with reduced learning to approach/avoid happy/angry individuals, respectively.

Conclusions Implicit spontaneous learning from socioemotional outcomes is not generally impaired in autism but may be specifically associated with autism severity in females with ASC, when they do not have an explicit strategy

---

†Rocco Mennella and Julie Grèzes contributed equally to this work.

*Correspondence:
Morgan Beaurenaut
beaurenaut.morgan@gmail.com
Julie Grèzes
julie.grezes@ens.psl.eu

Full list of author information is available at the end of the article

for adapting to their social environment. Clinical diagnosis and intervention ought to take into account individual differences in their full complexity, including the presence of co-morbid anxiety and depression, when dealing with social atypicalities in autism.

Keywords Autism, Social reinforcement learning, Emotional expressions, Approach/avoidance, Gender bias

Background

Autism spectrum condition (ASC or autism) is an early-onset, lifelong neurodevelopmental disorder characterized by symptom heterogeneity [1]. Individuals with ASC exhibit atypicalities in social communication and interactions, as well as repetitive and restricted behaviors and interests [2, 3]. In terms of social difficulties, individuals with ASC experience discomfort during eye contact and actively avoid it [4]. In addition, clinicians and family members report decreased attention to social cues and difficulties with reciprocity in social interactions, i.e., the ability to recognize and respond to socioemotional cues [5–8]. Consistent with clinical evidence, behavioral and neuroimaging findings reveal atypical perceptual processing of emotional and non-emotional social displays (e.g., faces and bodies) across development [9–16].

To explain these findings, the social motivation theory [17, 18] postulates that both atypical face processing and social difficulties in ASC are related to reduced sensitivity and responsiveness to social incentives. In other words, reduced motivation for social interaction - and associated hypoactivation of frontostriatal reward circuits - leads to reduced opportunities to learn and master social skills throughout development [see also 19]. Specifically, motivation refers to the anticipation of potentially rewarding and/or pleasurable stimuli ("wanting"), as well as the pleasure associated with reward consumption ("liking") and the ability to consider past rewards/punishments to guide future actions ("learning").

Recent studies support social motivation theory by showing that both structural and functional atypicalities in the mesolimbic reward pathway in response to social stimuli correlate with impairments in social interaction in individuals with ASC [20–23]. Nevertheless, a recent meta-analysis failed to find a strong relationship between diminished attention toward social stimuli and social seeking in ASC [24] and several criticisms have been raised to the social motivation theory [25]. For instance, it is unclear whether atypicalities in reward processing are specific for social situations. Recent studies have reported reduced reward wanting for both social and nonsocial reward, while no atypical reward liking emerged [26–31]. In this line, several meta-analyses confirmed that reward processing is atypical in autism across both social and non-social domains [32–34]. Furthermore, some studies found an unexpected neural and physiological hyperactivation of reward circuits in autism, suggesting that motivation is preserved but requires additional "neural effort", i.e. is less efficient in ASC [35–38].

Interestingly, and along with inconsistent results, previous research in ASC has focused mainly on reward anticipation and receipt, while reward learning has been less investigated. However, the ability to consider past rewards/punishments to guide future actions is crucial for optimally navigating our complex and dynamic social environment and is central to social motivation theory. Indeed, some studies support the idea that individuals with ASC have specific difficulties in learning from social rewards due to reduced sensitivity to social stimuli [32, 39–42], while other studies have found no differences in social learning between ASC and control groups [43].

Of note, learning deficits in social contexts may not only be related to reduced sensitivity to social stimuli, but also to more general difficulties in establishing stimulus-action-outcome associations in complex environments. Evidence suggests that ASC individuals have difficulty extracting the statistical structure, i.e., the predictability, of environments that vary in terms of action opportunities and their consequences [44–48]. This is the case in social settings that involve multiple actors in different interaction contexts, where the contingencies between actions and outcomes are dynamic and complex [49]. Accordingly, extracting action-outcome contingencies in social contexts has been shown to be altered in participants with ASC [50].

Overall, to effectively test social learning in autism, it seems essential to (1) provide sufficiently motivating social scenarios and (2) introduce complexity in the action-outcome contingency. Indeed, the heterogeneity of previous findings on social learning may be due to characteristics of the experimental paradigms used [51]. For example, it has been pointed out that the reward valence and magnitude of the social stimuli used in these studies - photos or videos of smiling faces of strangers or thumbs up/down - are difficult to quantify and may be insufficiently motivating [5, 28]. Furthermore, in classical paradigms, it is often quite explicit what the "good"/"bad" outcome is, regardless of its social nature. Thus, participants may simply understand the explicit rule that they should receive a thumbs-up or a smiling face, irrespective of whether they value it or not. This is particularly important because individuals with ASC often use more explicit learning strategies, sometimes leading to behavioral performance comparable with controls. For example, atypicalities in social cognition tend to disappear in

experimental paradigms with explicit instructions compared to implicit/spontaneous conditions [52, 53].

In addition to methodological considerations, the characteristics of the sample studied are also likely to play an important role in the heterogeneity of the results. First, reduced social motivation (social anhedonia) is observed transdiagnostically [54], not only in autism but also, for example, in depression [55] and anxiety [56, 57]. In particular, both reduced positive affect in healthy individuals and clinical anhedonia in depressed patients have been associated with reduced approach responses toward individuals with affiliative and prosocial expressions [55, 58]. This is relevant because compared to non-autistics, individuals with ASC have a higher lifetime prevalence of anxiety and depression of 42% and 37%, respectively [59], both of which predict lower social functioning in ASC [60, 61]. Furthermore, in both typically developing (TD) and ASC populations, female adults have higher rates of these co-occurring mental health conditions compared to males [62]. Of interest here, and in partial contradiction with this evidence, self-reported social motivation [63], as well as activation in the neural reward system, have been observed to be higher in females/girls relative to males/boys in both non-autistic and ASC individuals [64–66]. Overall, due to the great clinical heterogeneity, it is of paramount importance to rely on large sample sizes to account for inter-individual differences in ASC studies [67, 68].

The present study investigated social learning in ASC, by taking into account both the above-mentioned methodological aspects and the importance of inter-individual differences. We recruited a large online sample ($n=564$, 50% female, age- and education-matched) composed from participants declaring having ($n=274$) and not having ($n=290$) an ASC diagnosis. We relied on a previously validated paradigm [69] to test whether participants spontaneously (i.e., without instructions) develop a preference for the actions that lead to desirable socioemotional outcomes (i.e., avoiding threatening individuals/approaching friendly ones). Specifically, participants were presented with a waiting room containing four chairs, with the two chairs in the middle occupied by two individuals and the two outer chairs empty. They indicated which chair in the scene they would prefer to sit in by pressing a button, knowing that there were no right or wrong answers. At choice time, both seated individuals had a neutral expression, but after the response, one of the individuals changed the expression from neutral to emotional. Accordingly, in this brief feedback phase after choice, subjects would find themselves either far from (avoidance scenario) or close to (approach scenario) the individual expressing emotion. On each trial, one response (e.g., sit on the right chair) was associated with an 80% probability of avoidance and a 20% probability of

approach in the feedback phase. The other response (e.g., sit on the left chair) was associated with the complementary probabilities (20% avoidance, 80% approach). Crucially, the emotional individual was always present in the feedback phase, while the subject's physical distance from the emotional individual changed depending on response choice. Finally, we introduced variability to better simulate realistic social environments, by setting frequent action-outcome probability reversals. Here, we tested spontaneous learning using either angry or happy expressions, in different sub-samples.

Our main hypothesis was that participants with declared ASC would show diminished spontaneous social learning, as revealed by lower proportion of responses leading to desirable outcomes, i.e., avoidance of angry/approach of happy individuals, compared to controls. We also expected spontaneous learning of approach responses toward happy individuals to be more strongly impaired in ASC, based on the social motivation theory and on previous studies reporting reduced tendency to approach affiliative signals in ASC [70, 71]. We further explored (1) the influence of the emergence of explicit (reportable) approach/avoidance strategies [see 69]; (2) gender effects; (3) separate influence of autistic traits, depression and anxiety, and personality, as measured by questionnaires.

Methods

Participants
First, we run a version of the task in which, after participants' choice, one of the two individuals in the waiting room scene turned angry (angry feedback task, AFT; see *Reinforcement learning task*). We recruited initially $n=584$ participants, including a typical development group (TD; $n_{TD}=293$) and a group with autism (ASC; $n_{ASC}=291$). Participants were included in the TD group if they met the following pre-screening criteria: no diagnosis of mental disorder (including ASC), dementia, no visual impairment, no mental illness (even not diagnosed) having a significant impact on daily functioning, no need for psychiatric help, nor medication use. In addition, participants from Mennella et al. [69] sample were excluded from recruitment. Participants were included in the ASC group if they met the following pre-screening criteria: declared formal clinical diagnosis of autism spectrum disorder received from a psychiatrist, psychologist, or other qualified medical professional, no reported use of medication and no reported visual impairment. We then excluded participants meeting the following exclusion criteria: (1) TD individuals not matching in terms of age, gender, and education any ASC participant [MatchIt R package, matchit function, method exact −72]; (2) Reported issues with online feedback display (a-synchronous scene appearance and green tick not appearing on

Table 1 Participants' characteristics within each condition group (ASC and TD)

Variable	N	ASC, N=274[1]	TD, N=290[1]	p-value[2]	effect size[3]
Experiment	564			0.5	
Angry		127 (46%)	127 (44%)		
Happy		147 (54%)	163 (56%)		
Gender	564			0.5	
F		130 (47%)	146 (50%)		
M		144 (53%)	144 (50%)		
Strategy	564			0.4	
Non-explicit		185 (68%)	186 (64%)		
Explicit		89 (32%)	104 (36%)		
Age	564	24.0 (21.0, 28.0)	24.0 (21.0, 29.0)	> 0.9	0.004
Education	564			> 0.9	
Undergraduate		119 (43%)	130 (45%)		
A-levels		90 (33%)	92 (32%)		
Graduate		37 (14%)	41 (14%)		
Technical		25 (9.1%)	24 (8.3%)		
PhD		2 (0.7%)	2 (0.7%)		
Secondary		1 (0.4%)	1 (0.3%)		
Handedness	564			0.10	
R		231 (84%)	253 (87%)		
L		31 (11%)	33 (11%)		
A		12 (4.4%)	4 (1.4%)		
CATI total	564	154 (137, 168)	109 (90, 127)	< 0.001	0.664
CATIsoc	564	28 (23, 32)	19 (14, 24)	< 0.001	0.465
CATIcom	564	22 (18, 25)	13 (10, 17)	< 0.001	0.593
CATIcam	564	25 (21, 29)	18 (14, 22)	< 0.001	0.488
CATIrig	564	28.0 (24.0, 31.0)	22.0 (19.0, 26.0)	< 0.001	0.448
CATIrep	564	27 (23, 30)	19 (15, 23)	< 0.001	0.523
CATIsen	564	26 (21, 30)	16 (11, 19)	< 0.001	0.619
STAI-State	564	12.0 (9.0, 15.0)	10.0 (7.0, 12.0)	< 0.001	0.302
STAI-Trait	564	17.0 (13.0, 19.0)	12.0 (10.0, 15.0)	< 0.001	0.420
PHQ	564	11 (6, 16)	5 (2, 9)	< 0.001	0.408
BIS	564	24.0 (21.0, 26.8)	22.0 (19.0, 24.0)	< 0.001	0.245
BASDrive	564	10.00 (8.00, 12.00)	11.00 (9.00, 12.00)	0.4	0.036
BASFun	564	11.00 (9.00, 13.00)	12.00 (10.00, 13.00)	0.022	0.097
BASReward	564	16.00 (14.25, 18.00)	17.00 (15.00, 18.00)	0.035	0.089

[1]n (%); Median (IQR)

[2]Pearson's Chi-squared test; Wilcoxon rank sum test; Fisher's exact test

[3]Wilcoxon effect size, 0.10 - < 0.3 (small effect), 0.30 - < 0.5 (moderate effect) and >= 0.5 (large effect)

Note: Soc: Social Interactions, Com: Communication, Cam: Social Camouflage, Rig: Cognitive Rigidity, Rep: Repetitive Behaviour and Sen: Sensory Sensitivity

the chair; see *Reinforcement learning task*); (3) Task performance, namely if they responded to less than 90% of the trials in the allotted time, if they anticipated more than 15% of the total number of valid responses (the anticipation threshold was set at 100ms) or if they missed the 2 catch trials inserted into the auto-questionnaires. This resulted in a final sample of 254 participants to the AFT (n_{ASC}=127, see Table 1 for demographics).

Then, we ran a second version of the task in which after participants' choice, one of the two individuals in the waiting room scene turned happy (happy feedback task, HFT). Participants having been included in the AFT were not eligible for HFT. We initially recruited n=398 participants, including a TD group (n_{TD}=216) and an ASC group (n_{ASC}=182). We applied exactly the same inclusion/exclusion criteria as for the AFT, except the one concerning issues with online feedback display, as the task code was improved to avoid such issues. Furthermore, ASC group was recruited first, which allowed to ensure the best possible match with controls in the pre-screening phase. This resulted in a final sample of 310 participants who received happy feedback (n_{ASC}=147, see Table 1 for demographics). Participants' data from the AFT and the HFT were pooled, with the emotion in the feedback phase treated as a between-subjects variable. General socio-demographic characteristics of both samples and specific characteristics of the ASC sample are presented in Tables 1 and 2, respectively.

The experimental protocol was approved by the Comité d'Evaluation Ethique de l'Institut National de la Santé et de la Recherche Médicale (INSERM - IRB00003888 - N° 21–796) and was carried out in accordance with the Declaration of Helsinki. Participants were recruited online using the Prolific platform (https://www.prolific.com). They gave informed consent and were compensated for their participation in the study. The payment was £9 per hour (median experiment duration was 32.9 min).

General procedure

The experiment was coded in JavaScript, using the jsPsych 6.1.0 library [73] and was hosted on Pavlovia's servers (www.pavlovia.org). The experimental session consisted of four steps. First, participants provided their consent to participate, as well as sociodemographic information, and then completed the short version of the State Anxiety questionnaire [STAI-S-6; 74]. Second, after reading the instructions and becoming familiar with the task through a short training session, participants completed the reinforcement learning task. Third, at the end of the task, participants answered to several task-related questions. They reported whether they used a particular strategy to perform the task and whether they observed any changes in the faces. Participants who reported that they strategically tried to avoid angry individuals in the

Table 2 Characteristics of the ASC group

Declared ASC	N	F, $N=130$ [1]	M, $N=144$ [1]	*p*-value [2]	effect size [3]
Emotion	274			0.4	
Angry		64 (49%)	63 (44%)		
Happy		66 (51%)	81 (56%)		
Strategy	274			0.8	
Non-explicit		89 (68%)	96 (67%)		
Explicit		41 (32%)	48 (33%)		
Age	274	24.0 (21.0, 29.0)	23.0 (21.0, 28.0)	0.2	0.070
Diagnosis	274			< 0.001	
Autism spectrum disorder		75 (58%)	41 (28%)		
Asperger syndrome		30 (23%)	78 (54%)		
High functioning autism		18 (14%)	17 (12%)		
Low functioning autism		1 (0.8%)	3 (2.1%)		
Pervasive developmental disorder, not otherwise specified		2 (1.5%)	1 (0.7%)		
Other		4 (3.1%)	4 (2.8%)		
Diagnostic age	274	18 (12, 22)	12 (7, 18)	< 0.001	0.266
CATItotal	274	159 (139, 175)	148 (133, 164)	< 0.001	0.202
CATIsoc	274	29 (24, 32)	26 (21, 31)	0.027	0.134
CATIcom	274	23.0 (18.0, 26.0)	22.0 (17.8, 25.0)	0.3	0.067
CATIcam	274	27 (22, 30)	24 (20, 27)	< 0.001	0.242
CATIrig	274	28.0 (24.0, 32.0)	28.0 (24.0, 31.0)	0.6	0.029
CATIrep	274	28.0 (23.0, 31.0)	27.0 (22.0, 30.0)	0.14	0.090
CATIsen	274	28 (24, 32)	24 (19, 28)	< 0.001	0.287
Comorbidities					
Attention deficits	274	37 (28%)	42 (29%)	0.9	
Depression	274	74 (57%)	57 (40%)	0.004	
General Anxiety	274	65 (50%)	45 (31%)	0.002	
Social Anxiety	274	45 (35%)	39 (27%)	0.2	
Other Condition	274	39 (30%)	23 (16%)	0.006	

[1] n (%); Median (IQR)

[2] Pearson's Chi-squared test; Wilcoxon rank sum test; Fisher's exact test

[3] Wilcoxon effect size, 0.10 - < 0.3 (small effect), 0.30 - < 0.5 (moderate effect) and >= 0.5 (large effect)

Note: Soc: Social Interactions, Com: Communication, Cam: Social Camouflage, Rig: Cognitive Rigidity, Rep: Repetitive Behaviour and Sen: Sensory Sensitivity

AFT or to approach happy ones in the HFT were categorised into the Explicit strategy group, while the others were categorized into the Non-explicit strategy group [as in 69]. Fourth, participants carried out the subjective evaluation task, which began with the task instructions and a 1-minute training session. Finally, participants were debriefed about the subjective task as well as the overall

quality of the experimental session and ended the session by filling: (i) the short version of the Trait Anxiety questionnaire [STAI-T-6; 74] to assess individual variations in anxiety trait, (ii) the Comprehensive Autistic Trait Inventory [CATI; 75] constituted to six subscales that contribute independently to predicting autistic status (Social Interactions, Communication, Social Camouflage, Cognitive Rigidity, Repetitive Behaviour and Sensory Sensitivity), (iii) the Behavioural Inhibition-Activation Scales [BIS-BAS; 76] to assess individual tendencies to exhibit avoidance or approach behaviours via four subscales (BIS-inhibition, Bas-Drive, Bas-Fun Seeking and Bas Reward responsiveness) and (iv) the eight-item Patient Health Questionnaire depression scale [PHQ-8; 77] was used as a tool to evaluate current depression severity.

Reinforcement learning task

The stimuli depicted a waiting room with four chairs, with two individuals sitting in the two middle chairs [as in 78, 79]. Ten fixed pairs of same-gender identities (5 male and 5 female) were used [see 79 for full description and validation of the stimuli]. Each identity could be presented in both middle chairs, resulting in a set of 20 different stimuli. The fixed pairs of identities as well as the position of the identities in the scene were counterbalanced and each specific scene was presented 15 times in a random order (with the constraint that one stimulus could not be presented more than twice in a row). This resulted in 300 trials (2 genders × 5 pairs × 2 positions × 15 repetitions).

Participants were instructed that the experiment consisted of a series of decisions, and that they had to indicate where they would prefer to sit in the waiting room, while maintaining fixation on the fixation cross throughout the trial. Participants were asked to make spontaneous free choices and were informed that there were no correct answers. Each trial began with a grey screen displayed for 500ms, followed by a fixation cross superimposed on the upper centre of the screen for a duration varying between 500-700ms (see Fig. 1). Then, the scene with two individuals displaying neutral expressions appeared and remained on the screen until a valid response (consisting of a left or right button press within the given time) was registered, or until a maximum time of 1500ms, in the case of no response. If no response was registered, the message 'TOO SLOW' was displayed. To respond, participants used their keyboard, pressing "S" to sit in the left chair or "L" to sit in the right chair. Following the participants' response, a visual mask consisting of a random permutation of the pixels of the previously presented scene was displayed for 200ms, and feedback was provided immediately thereafter (500ms). This feedback presented the same scene, with a green tick symbol superimposed on the chosen chair, indicating the
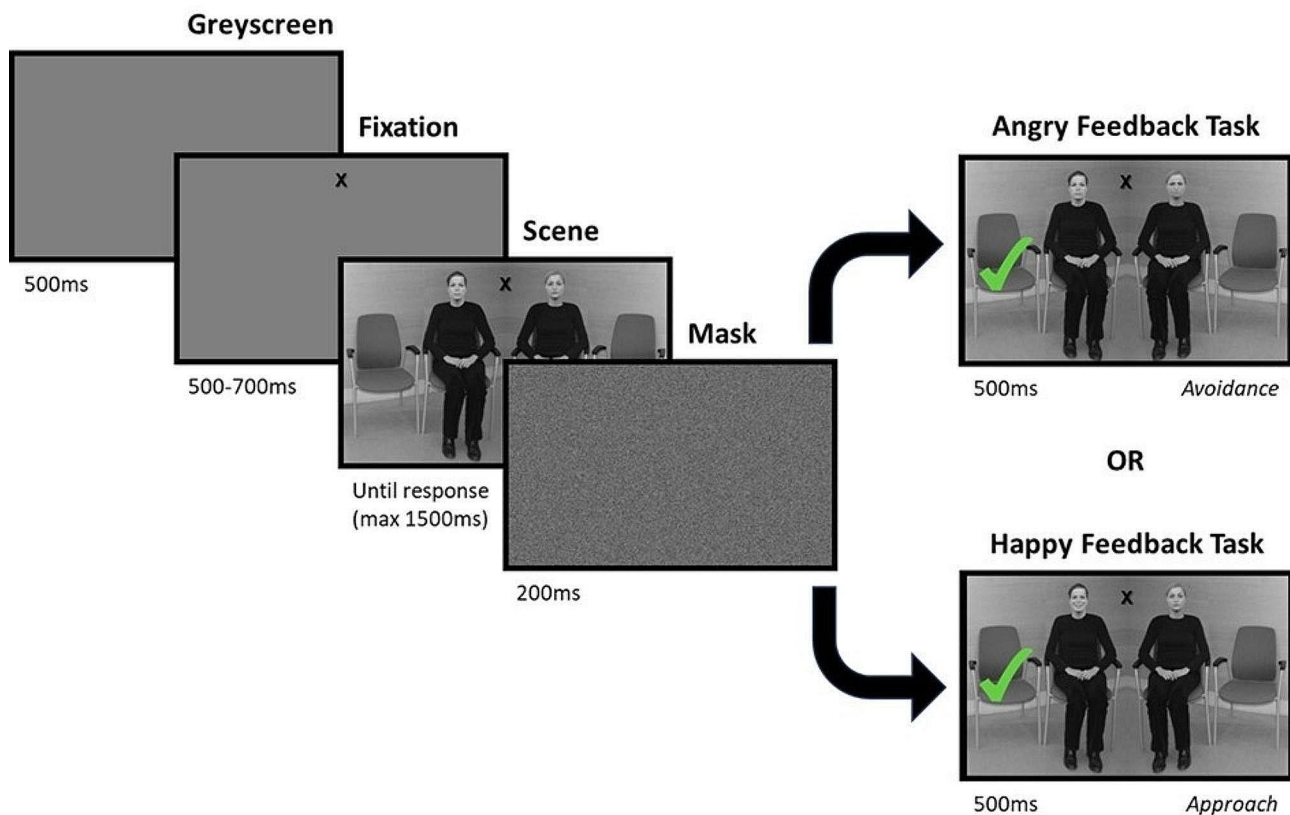
Fig. 1 Experimental design. Participants performed a reinforcement learning task in which they indicated their preference for a free chair by pressing a left/right button. During the feedback phase, they could find themselves either far from (avoidance) or close to (approach) an individual displaying an emotional expression (either happy for HFT or angry for AFT). The feedback was determined by a hidden probability associated with each chair. The figure illustrates an example trial in which the participant pressed the left button, resulting in the desired outcome for each version of the task: avoiding the angry individual or approaching the happy individual

participant's position in the scene. Importantly, either the individual near or far from the chosen chair changed their expression from neutral to angry in the AFT, or from neutral to happy in the HFT. Thus, the feedback to the participants was that they were seated far from (avoidance) or close to (approach) the emotional individual. On each trial, one response (e.g., sit in the right chair) was associated with an 80% probability of avoidance and a 20% probability of approach in the feedback phase. The other response (e.g., sit in the left chair) was associated with the complementary probability (20% avoidance, 80% approach). Importantly, the action-outcome probabilities were reversed on average every 25 trials (from 20 to 30), to ensure continuous goal-directed learning and to discourage habit formation.

Subjective evaluation task

During this task, participants had to indicate to what extent they would have liked to be in the situation represented by each of the 40 possible feedback which were presented in the Reinforcement Learning task: 10 pairs of actors, each of which could display either a neutral or an emotional expression, with a green tick indicating the participant's position either far from or close to the emotional actor. Each trial started with a grey screen displayed for 500ms, then a fixation cross was superimposed on the upper centre of the grey screen for a duration varying between 800 and 1200ms. Once the fixation cross disappeared (to allow for free visual exploration), the feedback scene was presented for 1000ms, followed by a visual analogue scale ranging from 'NOT AT ALL' to 'EXTREMELY' (re-coded offline as a continuous scale from 0 to 100). Participants were requested to provide their subjective evaluation on the visual analogue scale, by clicking on their cursor and moving it along the scale before validating their response by pressing 'continue'.

Statistical analyses: categorical approach

For the reinforcement learning, we treated as a first dependent variable the probability of "hits" (p(hits)), defined as the proportion of trials in which participants selected the seat associated with the higher probability of obtaining the desirable outcome (sitting far from the angry individual or sitting next to the happy individual).

In other words, p(hits) measures whether subjects spontaneously developed a preference for the action leading to a desirable socio-emotional outcome throughout the task (i.e., avoid angry and approach happy individuals). We studied the effect on p(hits) of four between-subject independent variables, namely Condition (TD vs. ASC), Emotion (Angry vs. Happy), Strategy (Explicit vs. Non-explicit) and Gender (Female vs. Male), by running a factorial analysis of variance (ANOVA).

To investigate sensitivity to the socio-emotional feedback, we analysed the proportion of 'repetition' (p(rep)), indicating whether participants repeated or changed the choice given at the previous trial (t-1), as a function of the received feedback (approach/avoidance). An ANOVA was conducted on p(rep) with the Feedback t-1 (Approach vs. Avoidance) as the within-subject factor and Condition, Emotion, Strategy and Gender as between-subject factors.

To complement the traditional frequentist statistical analyses, we quantified the relative strength of our empirical data in favor of the absence or presence of differences in learning between ASC and TD participants. To do so, a first Bayesian t-test was conducted on p(hits) with Condition as a between-subject factor. To conduct the second Bayesian t-test on p(rep), we first calculated, for each subject, the difference between the probability of repeating the previous choice when it led to a socially desirable outcome (approach joy / avoid anger) and the probability of repeating the previous choice when it led to an undesirable outcome (avoid joy / approach anger). We then performed the Bayesian t-test on this difference with Condition as a between-subject factor.

Statistical analyses: dimensional approach
In recent years, numerous studies have highlighted the limits of a categorical approach based on clinical diagnoses to psychopathology, proposing a shift to considering psychopathology in a dimensional and transdiagnostic manner [80–82]. Concerning autism, the use of DSM categories may be particularly problematic, due to the high rates of comorbidity and symptom variability [83, 84]. In this exploratory section, we therefore re-analysed our data adopting a dimensional approach. To do this, we applied a factor analysis to account for the partial overlap between the different subscale scores of our questionnaires and to identify potential latent structures between them. Factor analysis with Maximum likelihood Estimation was conducted using the *fa* function from the Psych R package [85], with an oblique rotation (varimax). Thirteen subscales were entered as measured variables in the factor analysis (the 6 subscales of the CATI questionnaire, the STAI-state and STAI-trait, the PHQ questionnaire, the 4 subscales of the BIS/BAS). We selected the number of factors using the function *n_factors* from the

R package Parameters [86] that automatically estimates the optimal number of dimensions to retain, crossing several methods. We then tested the extent to which each of the identified factors interacted with our between-subjects variables (Condition, Emotion, Strategy, Gender) on p(hits). Finally, we ran an ANOVA on the identified factors with Condition, Emotion, Strategy, Gender as between-subjects variables.

Statistical analyses: Subjective evaluation of the feedback
For the subjective evaluation task, we treated as a dependent variable the subjective rating attributed to each feedback. An ANOVA was conducted on the subjective ratings with the nature of the Feedback (Approach vs. Avoidance) as within-subject factor and Condition, Emotion, Strategy and Gender as between-subject factors. We also investigated the relationship between participants' subjective evaluations of angry and happy feedback and socioemotional learning (e.g., p(hits)). To do so, we computed the difference in the subjective ratings between avoidance and approach feedback (diffAvAp), and tested (ANCOVA) the extent to which this difference in subjective ratings interacted with our between-subject variable Emotion on p(hits). All the analyses have been performed on R [87] via Rstudio [88] and JASP [89].

Tables of descriptive and inferential statistics of all the analyses described above (Tables S1-17, Figures S1-3), as well as of additional control analyses (on possible confounding variables (Tables S18-21, Figures S4-6), on the role of stimuli's gender on participants' performance (Tables S22-24, Figure S7) and on the relationship between age of diagnosis and autism severity (Figure S8)) are available in the Supplementary Materials.

Results
Spontaneous learning of action-outcomes contingencies
A main effect of Emotion on the probability of choosing the action that maximises the chance to obtain desirable socio-emotional outcome (p(hits)) emerged, $F(1,548) = 12.6$, $p < 0.001$, $\eta^2_p = 0.023$, indicating stronger spontaneous learning to avoid angry individuals (AFT), compared to approaching happy ones (HFT). A main effect of Gender [$F(1,548) = 24.9$, $p < 0.001$, $\eta^2_p = 0.043$] and a main effect of Strategy [$F(1,548) = 60.2$, $p < 0.001$, $\eta^2_p = 0.099$] also emerged, further characterised by an interaction between Gender and Strategy [$F(1,548) = 4.8$, $p = 0.029$, $\eta^2_p = 0.009$], indicating that female participants with an explicit strategy chose the rewarding chair more often than all other participants (see Fig. 2A and B). Interestingly, all groups characterizing the interaction between Gender and Strategy showed learning significantly above chance level (all p-values for one sample t-tests against 0.5 < 0.001). No other main effects or interactions, notably with the clinical Condition, were found
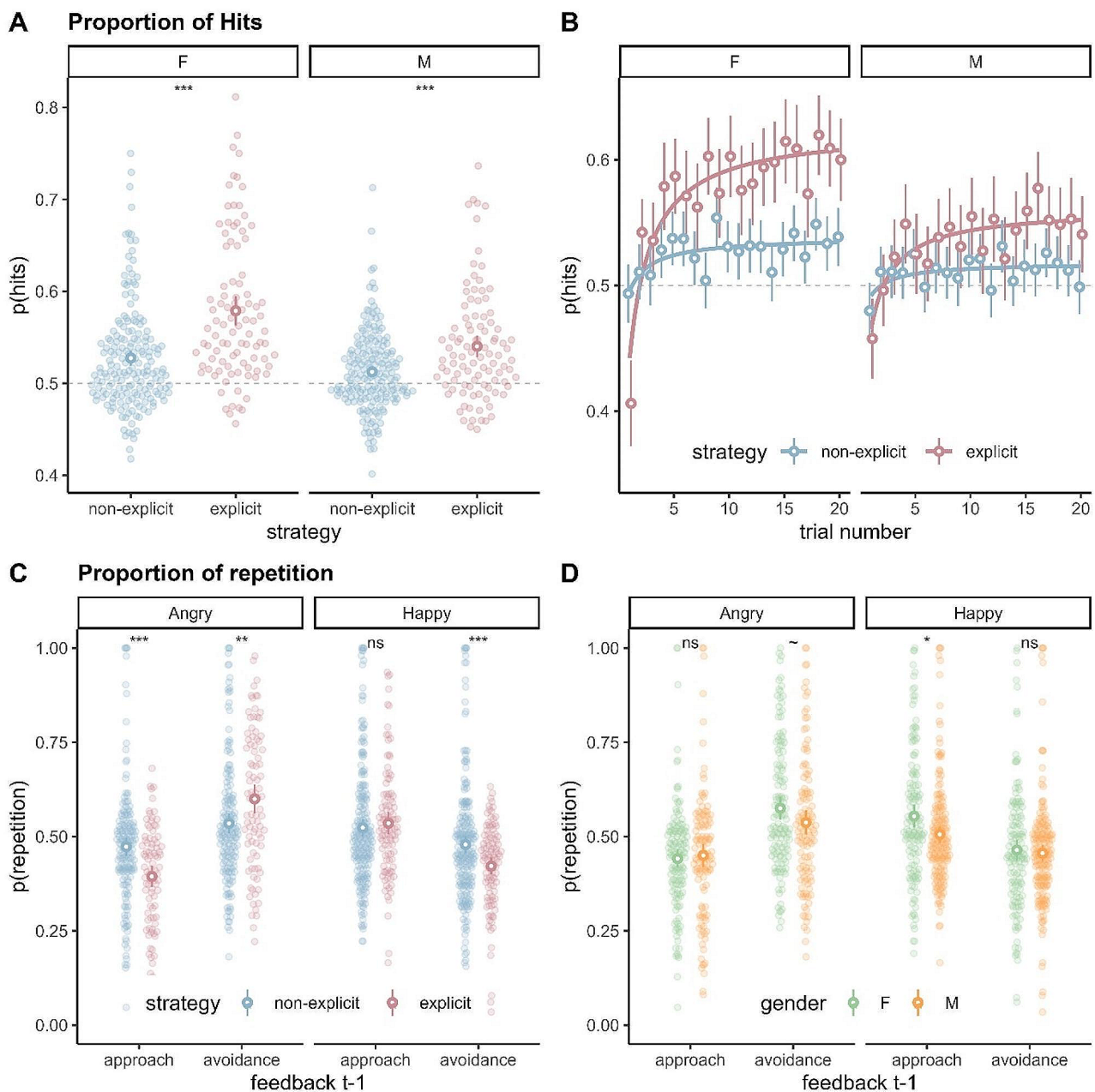
Fig. 2 Summary of behavioural results from the categorical approach. Thick-contour dots represent means, while error bars represent 95% confidence intervals for the normal distribution. Shaded points represent individual participant means. A: Mean proportion of hits across the task, as a function of participant strategy and gender. B: Mean proportion of hits over the first 20 trials across blocks of stable action-outcome contingencies (trial 1 = reversal trial). The fitted curve represents the best fit (and 95% confidence interval) for the hyperbolic function $y = 1 - 1/(1 + x)$ previously used in Mennella et al. (2022) and confirms the results for the mean proportion of hits (A). **C & D**: Mean proportion of action repetitions following either approach or avoidance outcomes in the angry and happy versions of the reinforcement learning task, as a function of participant strategy (C) and gender (D). Symbols: $p < 0.001$ ***; $p < 0.01$ **, $p < 0.05$ *, $p < 0.1$ ~

to be significant (all ps > 0.054, see Tables S1-2). Furthermore, Bayesian t-test with Condition as between-subject factor revealed a moderate level of evidence supporting similar level of spontaneous learning of action-outcomes contingencies between ASC and TD (BF$_{10}$=0.11, see Figure S1).

Adaptation to immediate socio-emotional feedback

The ANOVA revealed a significant Gender x Condition interaction [F(1,548)=4.8, p=0.029, η^2_p = 0.009], indicating that female compared to male participants repeated their previous choice more often, regardless of its outcome, in the ASC group but not in the

TD group. We also observed a main effect of the Feedback t-1 [$F(1,548) = 15.0$, $p < 0.001$, $\eta^2_p = 0.027$], significant interactions between Feedback t-1 x Emotion [$F(1,548) = 281.2$, $p < 0.001$, $\eta^2_p = 0.34$], and Feedback t-1 x Strategy [$F(1,548) = 7.6$, $p = 0.006$, $\eta^2_p = 0.014$], further characterised by a three-way interaction between Feedback t-1 x Strategy x Emotion [$F(1,548) = 69.7$, $p < 0.001$, $\eta^2_p = 0.11$]. Participants more often repeated previous choices that led to desirable social outcomes (avoid anger & approach happy) compared to undesirable social outcomes (approach anger & avoid happy); this difference in p(rep) between desirable and undesirable outcomes was enhanced in participants in the Explicit group, particularly in the Angry task (see Fig. 2C). In other words, participants with an explicit strategy had a higher sensitivity to avoidance of angry individuals than to approach of happy ones.

Finally, the interaction Feedback t-1 x Gender x Emotion interaction was significant [$F(1,548) = 15.8$, $p < 0.001$, $\eta^2_p = 0.028$], such that females more often repeated previous choices that led to desirable outcomes (avoid anger & approach happy) compared to males, but not to undesirable outcomes (see Fig. 2D). In the above-mentioned analysis, we observed that females performed better on the task compared to males. To assess whether males were nonetheless sensitive to social feed-back, we conducted two additional ANOVAs. The first, which focused on all males, revealed that the three-way interaction between Feedback t-1 x Strategy x Emotion was significant [$F(1,280) = 30.6$, $p < 0.001$, $\eta^2_p = 0.010$]. The second, focusing only on males in the Non-explicit group, also revealed that the interaction between Feedback t-1 x Emotion was significant [$F(1,185) = 26.87$, $p < 0.001$, $\eta^2_p = 0.13$]. No other main effects or interactions, notably with the Condition, were found to be significant (all ps > 0.12, see Tables S3-6). Consistently, Bayesian t-test with Condition as between-subject factor revealed a strong level of evidence in favor of similar level of adaptation to immediate socio-emotional feedback between ASC and TD ($BF_{10} = 0.095$, see Figure S2).

Dimensional approach of spontaneous learning
First, we tested the factorability of our data using Kaiser-Meyer-Olkin (KMO) (overall MSA = 0.87, suggested cutoff KMO > 0.6), and Barthlett's test of sphericity ($\chi2(78) = 3.492$, $p < 0.001$), both of which confirming the usefulness of performing a factor analysis on our data. A 3-factor latent structure was found to explain the shared variance best and most parsimoniously between all 13 subscale scores (supported by 9 (33.33%) out of 27 methods). Factor1, labelled 'Autistic traits', regrouped all the subscales of the CATI questionnaire and explained 26% of the variance. Factor2, labelled 'Anxiety-Depression' (17% of the variance), regrouped Trait and State anxiety

(STAI), PHQ-depression, and the BIS. Finally, Factor3, labelled 'Appetitive motivation' (12% of the variance), included the 3 subscales of the behavioural activation system (BAS).

The ANCOVA (see Tables S7-10) revealed a significant three-way interaction with Anxiety-Depression [Factor2*Gender*Strategy, $F(1,500) = 5.199$, $p = 0.023$, $\eta2p = 0.010$]. Correlational analyses indicated that the interaction emerged due to a significant negative correlation in the Non-explicit female group between the anxious-depression factor and p(hits) ($r = -0.16$, $p = 0.035$, see Fig. 3B), while correlations were non-significant in the other groups. Furthermore, the ANCOVA revealed a significant interaction between Condition, Emotion, Strategy, Gender and Factor1-Autistic traits [$F(1,500) = 10.189$, $p = 0.002$, $\eta2p = 0.02$]. To better characterise this interaction, we ran an additional ANCOVA, separately for ASC and TD, including only the Autistic trait factor. The ANOVA on the ASC, but not on the TD, revealed a significant four-way interaction between Emotion, Strategy, Gender and Autistic traits [$F(1,258) = 9.636$, $p = 0.002$, $\eta2p = 0.036$], further characterised by a significant interaction between Emotion, Gender and Autistic traits in Non-explicit participants only [$F(1,177) = 8.825$, $p = 0.003$, $\eta2p = 0.048$]. This latter interaction is illustrated by a negative correlation in ASC female participants from the Non-explicit group between p(hits) and Autistic traits ($r = -0.4$, $p = 0.0048$, Fig. 3A).

The ANOVA on Factor1-Autistic traits revealed a significant interaction between Condition and Gender [$F(1,548) = 16.233$, $p < 0.001$, $\eta2p = 0.029$, see Fig. 3C and Table S11], characterised by higher levels of autistic traits in females compared to males in the ASC group, and the reverse in the TD group. For the Factor2-Anxiety-Depression, a significant interaction between Condition, Gender and Emotion emerged [$F(1,548) = 4.218$, $p = 0.040$, $\eta2p = 0.008$, see Fig. 3D and Table S12], characterised by higher levels of Anxiety-Depression in females compared to males, except in the Angry feedback task for the ASC group.

Subjective evaluations of angry and happy feedback
The ANOVA revealed an interaction between Condition, Strategy, Emotion, Gender and Feedback [$F(1,548) = 4.17$, $p = 0.041$, $\eta2p = 0.008$] (Tables S13-16). To better characterise this interaction, we conducted an additional ANOVA, separately for Angry and Happy tasks. The ANOVA on the AFT, but not on the HFT, revealed two interactions with Condition. First, a significant three-way interaction between Condition, Gender and Feedback [$F(1,246) = 4.98$, $p = 0.026$, $\eta2p = 0.020$] (Fig. 4A) was related to the fact that ASC females rated the avoidance scenario less positively than ASC males, whereas TD females rated the approach scenario more
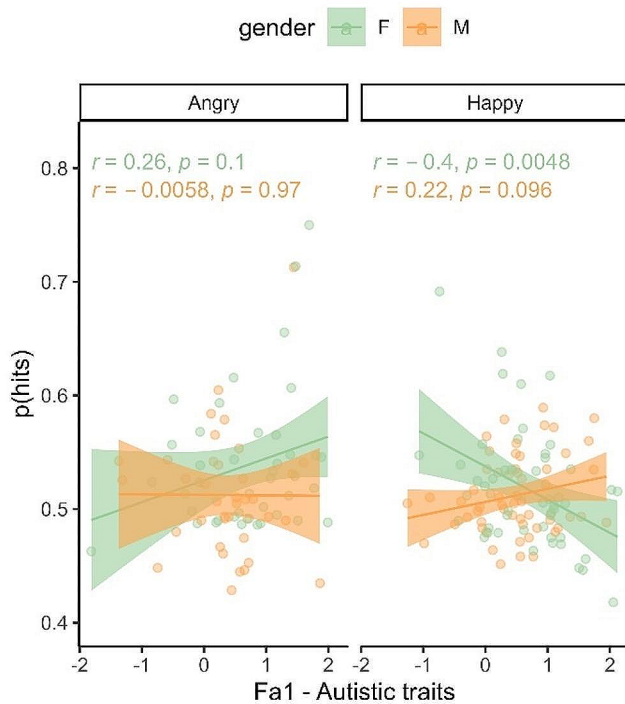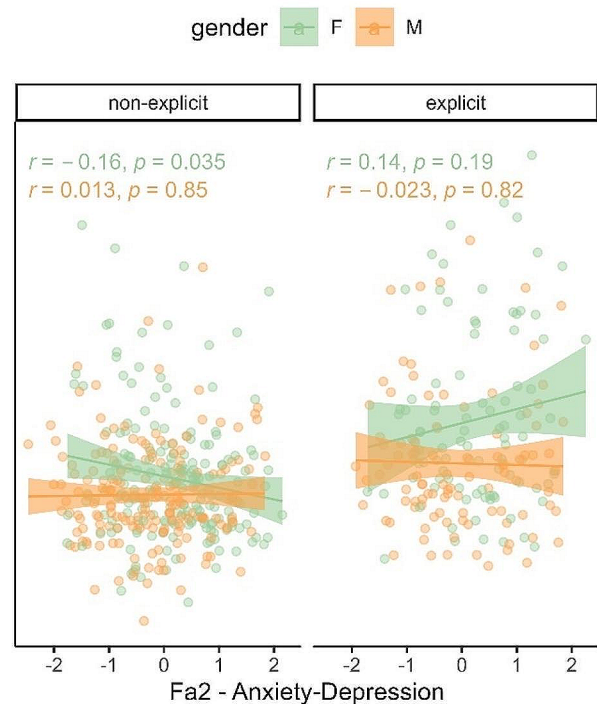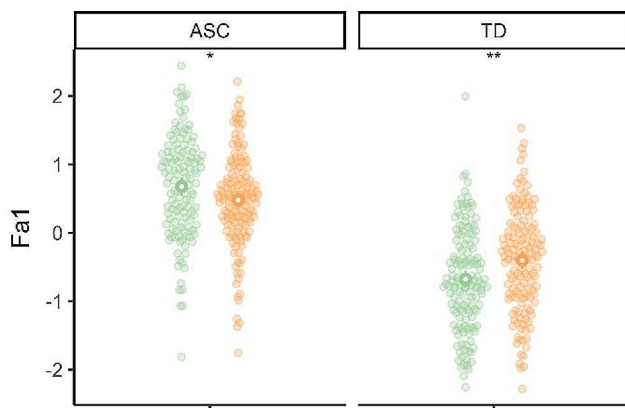
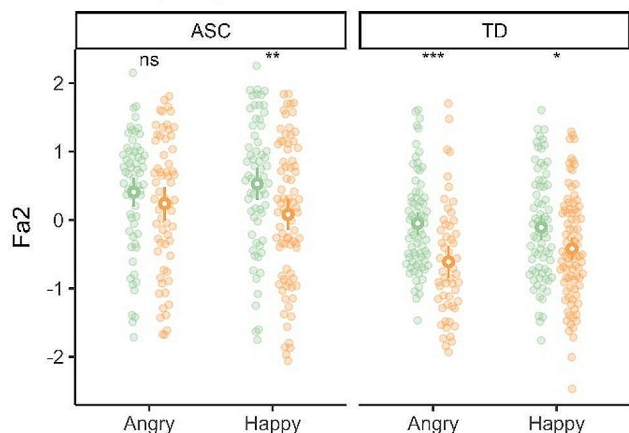**Fig. 3** Summary of behavioural results from the dimensional approach. A: For ASC participants who did not develop an explicit learning strategy, Pearson correlations between mean proportion of hits and the Fa1-Autistic traits, as a function of Angry or Happy feedback and participant gender. **B**: In the Happy feedback task, Pearson correlations between the mean proportion of hits and the Fa2-Anxiety-Depression score, as a function of participant strategy and gender. **C**: Mean Fa1-Autistic traits as a function of participant condition and gender. **D**: Mean Fa2-Anxiety-Depression score in the angry and happy versions of the reinforcement learning task, as a function of participant's condition and gender. Same convention as Fig. 2

negatively than TD males. The second significant three-way interaction between Condition, Strategy and Feedback [F(1,246)=4.83, *p*=0.029, η2p=0.019] (Fig. 4B) was related to the fact that the Explicit participants rated the approach scenario more negatively than the Non-explicit participants, and this difference was more important in ASC compared to TD. In contrast, the ANOVA on the HFT task only revealed a two-way interaction between

Strategy and Feedback [F(1,302)=30.98 *p*<0.001, η2p=0.093] (Fig. 4C), characterized by more positive ratings of approach feedback and more negative ratings of avoidance feedback from the Explicit participants compared to the Non-explicit participants.

Finally, we explored the relationship between participants' subjective evaluations of angry and happy feedback and socioemotional learning. The ANCOVA
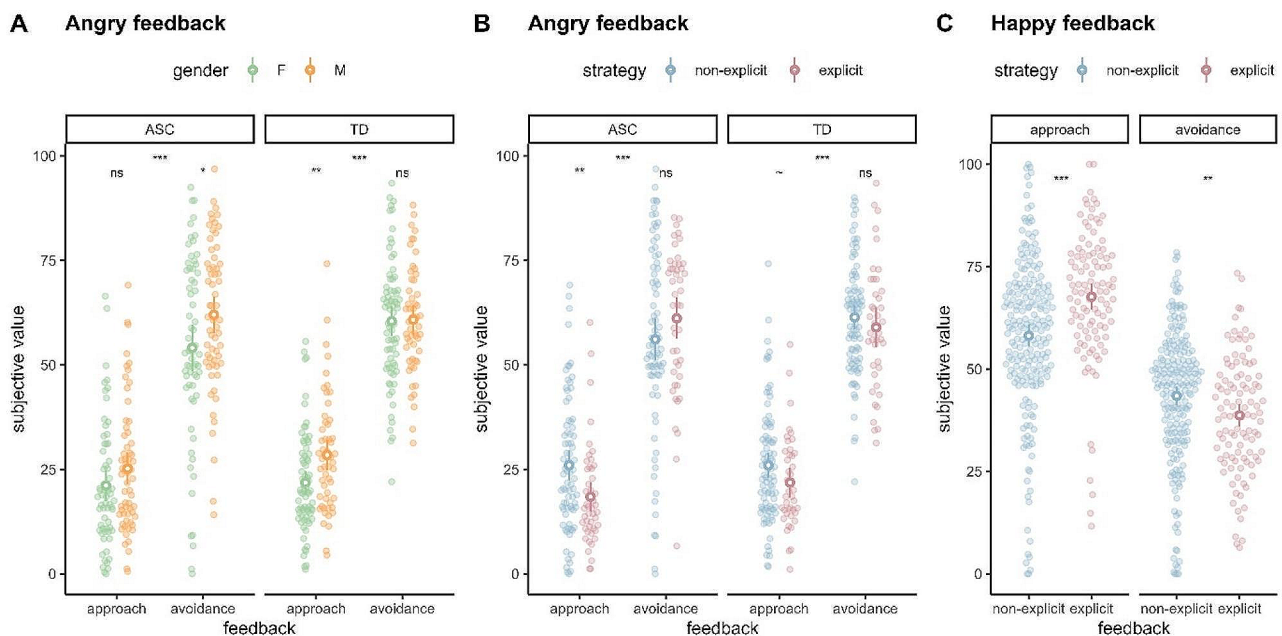
Fig. 4 Results of the subjective evaluation task. Thick-contour dots represent means, while error bars represent 95% confidence intervals for the normal distribution. Shaded points represent individual participant means. **A**: Mean subjective evaluation of approach and avoidance feedback in the Angry feedback task, as a function of participant condition and gender. **B**: Mean subjective evaluation of approach and avoidance feedback in the Angry feedback task, as a function of participant strategy and condition. **C**: Mean subjective evaluation of approach and avoidance feedback in the Happy feedback task, as a function of participant strategy. Same convention as Fig. 2

revealed a significant two-way interaction between Emotion and diffAvAp (i.e. the difference in the subjective ratings between avoidance and approach feedback; [$F(1,560)=35,23$, $p<0.001$, $\eta2p=0.059$], see Table S17). Correlational analyses indicated that the interaction was due to a positive correlation with learning (p(hits)) in the angry feedback task ($r=0.21$, $p<0.001$) and a negative correlation with learning in the happy feedback task ($r = -0.29$, $p<0.001$, see Figure S3).

Discussion

The present study set out to investigate whether socio-emotional learning in autism is atypical, as predicted by the social motivation theory [17, 18]. To this end, we tested whether participants spontaneously learned from socio-emotional outcomes, in a social reinforcement learning task. Importantly, to target implicit learning, participants were neither informed of the presence of emotional displays, nor instructed to learn. To account for the heterogeneity of autism, we ran an online task on a large cohort of TD and ASC participants, matched for gender, age, and education. Contrary to our main hypotheses, all participants, regardless of their diagnostic group, spontaneously learned to choose the action that led to the most socially desirable outcome, i.e., to avoid angry individuals and approach happy ones. In both groups, approximately one-third of participants developed an explicit angry-avoidance/happy-approach strategy and

showed better learning than the non-explicit group, replicating previous results in typically developed participants [69]. When accounting for dimensional variations in autistic traits, depression and anxiety, and personality, two main findings emerged, specifically in females who did not report an explicit learning strategy: within this subgroup, (1) higher levels of autistic traits in ASC participants correlated with reduced spontaneous learning to approach happy individuals; (2) higher levels of anxiety-depression across both ASC and TD participants correlated with reduced spontaneous learning to approach/avoid happy/angry individuals, respectively.

One of the main hypotheses of the social motivation theory of autism is that reduced social motivation leads to reduced spontaneous learning and mastering of social skills throughout development [17, 19]. Contrary to this prediction, the first important finding of the present study is that both frequentist and Bayesian analyses converged on the absence of clear differences in socio-emotional learning between ASC and TD participants in our large and gender-balanced adult samples. Of note, what motivated learning in the present task was the physical distance in the waiting room between the subject and the individual expressing the emotion, which is intrinsically social. Accordingly, at the subjective level, both TD and ASC reported a preference for approaching happy/avoiding angry individuals, compared to avoiding happy/approaching angry individuals. Moreover, as

indicated by the results on the proportion of response repetition, both groups considered the previous feedback (approach/avoidance) to adapt their subsequent responses. Thus, overall ASC participants did not show reduced sensitivity to socio-emotional feedback compared to TD, either at the subjective or behavioral level. In addition, it has been argued that atypicalities in learning in social contexts might derive from a more general difficulty in extracting statistical regularities from complex environments [44–48]. Here, to reproduce the complexity of social environments, the association between action and social outcome was probabilistic and multiple reversals of such contingency were introduced. Still, ASC participants were able to show statistically significant learning, comparable to the TD group, as indicated by the proportion of hits across trials and within blocks of stable action-outcome contingencies. Finally, on average, females showed higher spontaneous learning than males, regardless of clinical condition, emotion, and strategy. This is consistent with previous evidence from the literature showing that self-reported social motivation [63], social interaction and social communication skills, as well as activation in the neural reward system, are higher in females/girls compared to males/boys in both non-autistic and ASC individuals [64–66].

Importantly, in the present study, we adopted an implicit learning task that provided no learning instructions and never mentioned the presence or relevance of emotional expressions. Accordingly, most participants, both ASC and TD, did not report using an explicit approach/avoidance strategy at the end of the task. Nevertheless, they still showed sensitivity to the feedback and significant learning on average [see 69]. In the past, learning in ASC individuals has been found to be typical in both implicit non-social contexts [90, 91] and explicit social contexts [e.g., 26]. Our study further extends these results by showing that implicit learning abilities can be preserved in social situations in ASC. On the other hand, approximately one-third of participants reported spontaneously developing an explicit approach/avoidance strategy during the task, replicating previous findings on TD individuals [69]. The explicit learning strategy group showed better learning than the non-explicit group, and this effect was more pronounced in females than in males. As in the non-explicit groups, no differences between ASC and TD emerged in the explicit groups, suggesting that ASC individuals can form explicit behavioral strategies for responding in social contexts. However, more research is needed to determine whether individuals who can form explicit behavioral strategies are better at using them effectively to achieve desirable socio-emotional outcomes, or whether, conversely, the strategy is more likely to be consciously understood in individuals who have a stronger learning effect.

Interestingly, dimensional analyses of autistic traits, depression and anxiety, and personality, revealed that these traits were correlated with learning only in a specific subgroup, namely the female participants who did not develop an explicit approach/avoidance strategy (hereafter, female non-explicit; F-NE). First, within the F-NE participants with ASC, the severity of the self-reported autistic traits was negatively correlated with the ability to learn the response that maximized the chances of approaching happy individuals. In other words, F-NE ASC participants showed reduced learning from affiliative social feedback, proportionally to the reported severity of their ASC symptoms. This finding is partially consistent with our main hypothesis of stronger impairments in learning from affiliative social signals in ASC, based on the social motivation theory and on previous studies reporting reduced tendency to approach affiliative signals in ASC [70, 71]. Yet, this finding is intriguing because not only did females show higher spontaneous learning than males, on average, but females with ASC also reported more severe autistic symptoms (CATI total), compared to males. How can females with ASC have both more severe autistic symptoms than males and better performance on the social learning task?

In a recent study using the CATI, self-reported autistic traits were found to be higher in females than in males with ASC, and the opposite was found in the TD population (personal communication from Dr. Michael English, November 2023), similar to the present study. Autistic adults without intellectual disability, especially females, are known to use compensatory strategies to mask autistic characteristics and camouflage their social difficulties [92–95]. Camouflage can be measured using either self-report or internal-external discrepancy approaches [96]. Accordingly, a first explanation for the higher autistic traits in ASC females is that the CATI questionnaire includes a camouflage subscale, on which females with ASC did score higher than males. However, the same pattern of results has also been found in some previous studies using other self-report scales that do not systematically account for camouflage, such as the Autism spectrum Quotient [AQ; 97–99] and the Broad Autism Phenotype Questionnaire [BAPQ; 96].

Therefore, a second explanation for our findings relates to the fact that camouflage can be operationalized as the discrepancy between external behavioral presentation in social contexts and internal status - self-representation of autistic symptoms [94]. To conform to societal expectations and pressures regarding gender role's stereotypes [e.g., 100,101], females with ASC may be more preoccupied with their socio-emotional abilities (self-reflection). As a result, they may have a better representation of their symptoms, hence their higher scores on the autistic questionnaires mentioned above. Conversely, when

measures of socio-emotional abilities are based on clinical behavioral assessments rather than self-report (e.g., with the ADOS and ADI), ASC females, who more often camouflage their social difficulties, tend to have either lower [94, 102] or similar autistic scores [103] than males. Therefore, self-report of autistic symptoms may currently be poorly comparable across genders, and this may also be the case for clinical evaluation, but with a gender bias going in the opposite direction. More research is needed to determine whether gender differences in self-report, as well as in clinical evaluation, correspond to differences in the severity of the social atypicalities.

The fact that the predictions of the social motivation theory seemed to hold specifically for the F-NE ASC subgroup, as indicated by the correlation between autism severity and learning, may be related to two concurrent aspects. On the one hand, socio-emotional difficulties in ASC females are more predictive of an autism diagnosis than in boys/males, whereas the opposite is true for repetitive and restricted behavior symptoms [104, 105]. Therefore, there may be a closer correspondence between variations in CATI scores and behavioral difficulties in social contexts within the female subgroup. On the other hand, this relationship may be stronger for the F-NE ASC participants, precisely because they failed to develop an explicit strategy for approaching happy individuals. The reliance on explicit compensation strategies and social scripts to behave adaptively in social contexts is particularly pronounced in females with ASC [106, 107], often leading to under- or mis-estimation of their social difficulties [102, 108] and later diagnosis compared to boys/males [e.g., 109–111]. Here, females with ASC who did not extract and make explicit the task's regularities may have failed to compensate for their reduced sensitivity and responsiveness to affiliative signals, explaining the clearer relationship between ASC severity and learning in this subgroup.

Our dimensional analyses further revealed a correlation between anxiety/depressive symptoms and learning performance in the F-NE subgroup, across conditions. It is well known that the prevalence of anxious and depressive symptoms is higher in females than in males in both TD and ASC populations [62] and this was also true in the present study. Furthermore, both anxiety and depression have been associated with difficulties in learning from reward [112,113, for a review, see 114], decision making [for a review, see 115], socioemotional behavior [55, 116–118] and predict lower social functioning in ASC [60, 61]. Therefore, female participants who did not develop an explicit approach/avoidance strategy were more likely to manifest this relationship. This highlights the importance of adopting a transdiagnostic perspective on ASC, taking into account the co-occurring clinical symptoms when assessing socio-emotional behavior, in order to capture the full complexity of the social atypicality. Accordingly, compared to TD, individuals with ASC have a high lifetime prevalence of anxiety and depression of 42% and 37%, respectively [59], and a similar pattern of findings was found here (see Fig. 3D; Tables 1 and 2 & S12). In summary, depressive and anxiety symptoms might concur with autistic traits in determining each individual's social functioning, and ought to be considered in both research and clinical assessment, especially in interaction with gender.

Finally, we examined the role of possible confounding variables that may have influenced learning, beyond the effect of sensitivity to social reward. First, we focused on the ability to remain attentive throughout the task. On the one hand, it is conceivable that learning might deteriorate over the time of the task, which lasted around 30 min, due to a decrease in attention. On the other hand, self-reported comorbidity with attention deficit disorder in the ASC group might also have negatively impacted performance. Nonetheless, in our sample, neither the effect of time (splitting the experiment into 3 parts), neither the presence of self-reported diagnosis of attention deficit, showed a significant effect on participants' learning performance (See Supplementary Materials Tables S19-S20 and Figure S5).

Additionally, although there is conflicting evidence regarding altered emotional face processing and recognition in autism [119, 120], these atypicalities may also affect performance. To account for interindividual differences in emotion recognition abilities, we asked each participant during the debriefing phase whether they had seen emotions on the faces, allowing us to distinguish participants who had seen emotions from those who had not. We observed that while there was no difference between the diagnostic groups in the angry feedback task, in the happy task, ASC participants reported seeing the emotion less often compared to TD (See Supplementary Materials Table S21 and Figure S6). However, not only did reporting having seen the emotion not correlate with better learning in the happy feedback task, but our main finding in the present paper is that there is no significant difference in learning between diagnostic groups. Thus, despite ASC participants differing in the ability to explicitly recognize and report affiliative social signals, this effect is independent from their ability to use this information to adapt their behavior.

Limitations

This study should be interpreted in light of some limitations. First, to obtain a large and balanced sample, we recruited participants online, and relied on reports of formal clinical diagnosis. Recently, it has been highlighted that some participants may misrepresent themselves as autistic online for financial compensation [121].

Nonetheless, on the Prolific platform, participants who identify as having ASC are automatically excluded from studies focusing on TD participants. Given that experimental studies on TD individuals are more common, such a strategy is financially disadvantageous. Moreover, it is important to emphasize that we asked participants whether they had received a formal diagnosis from a clinician, the age at which this diagnosis was made, the specific form of autism, their comorbidities (both explicit and self-reported), and their self-reported autistic traits. This approach allowed us to confirm that our recruitment matched the known characteristics of the ASC population compared to TD, namely high levels of autistic traits, higher presence of comorbidities - especially anxiety and depression - and late diagnosis for women compared to men with ASC [84, 109, 122].

Another limitation is that, due to the nature of our online task, we selected mostly high-functioning adults with ASC, which restraints the generalizability of our conclusions. On the one hand, this may be relevant in relation to the social motivation theory, which postulates that learning in social contexts is impaired throughout development, beginning with early life experiences. On the other hand, learning may be more severely impaired in individuals with intellectual disabilities. Nonetheless, our conclusions still suggest that impaired social learning is unlikely to be a common and general mechanism underlying social difficulties in ASC, as suggested by the social motivation theory. While our result by itself contradicts the main claim of the social motivation theory, namely that motivation toward social reward is diminished in ASC, methodologically it would have been important to include a non-social learning task, as this could provide a full picture of learning from different kinds of reward in ASC. In the context of the present online study, it would have been difficult timewise to demand subjects to perform an additional nonsocial task. Instead, we preferred to include a second social reinforcement learning task with happy individuals to generalize results across negatively and positively valenced expressions using a between-subject experimental design. Also, future studies ought to test implicit socioemotional learning in children and/or adolescents with ASC. Indeed, we cannot exclude the possibility that our adult sample had learned how to improve their social skills, either through life experiences or cognitive and/or behavioral training (such as Applied Behavior Analysis; ABA, [123]). Partially in line with this hypothesis, we showed that the longer the time since diagnosis, the less severe the self-reported autistic traits (first factor from the factor analysis; see Supplementary Materials, Figure S8). Of note, this may be relevant to gender differences as, in line with the literature, female ASC were diagnosed later than males [111, 124]. The lack of information on the previous exposure of our sample to behavioral/cognitive interventions, which could indeed reduce social difficulties, is a limitation of the present study.

Furthermore, the lack of neural measures in the present study prevents us from drawing strong conclusions about the underlying neural mechanisms that are central to the social motivation theory. This is important because many previous studies have failed to find behavioral differences in social tasks, but have nevertheless found neural functional differences, either in the form of generally reduced [28, 125], or generally increased activity [35, 36, 126], or reduced/increased functional connectivity [127] between brain regions. Future neuroimaging studies with large and gender-balanced samples could help to further clarify the specificities in social learning that emerged in female participants without an explicit strategy in the present study. For example, we hypothesize that spontaneous learning in this subgroup would correlate with the activation of the fronto-striatal dopaminergic system, whereas learning via an explicit strategy might depend on regions related to theory of mind (superior temporal sulcus, temporal-parietal junction and dorsal anterior cingulate) or cognitive control (dorsolateral prefrontal cortex).

Finally, the socio-emotional feedback provided in the present task may have been too fast and subtle to allow for significant learning in the majority of the population. At the group level, both TD and ASC showed significant adaptation to feedback and learning, whereas within-individual learning was only significant for a minority of participants. This was particularly true for males, which may have prevented the emergence of more pronounced differences between ASC and TD. While spontaneous and implicit social learning is a crucial building block for the development of everyday life social skills, it will be interesting in future studies to compare results with tasks that allow for an easier extraction of action-outcome probabilities of the social environment. Lastly, participants here were forced to choose between two social actions. However, the differences between TD and ASC may be more pronounced in a setting that mimics real-world conditions in which a non-social action is available (e.g., refraining to sit). Indeed, some ASC participants reported during debriefing that they "personally would not sit next to anyone anyway," especially if individuals "might want to start a conversation," and expressed a preference for a missing "stand outside option" that they "would press that quite often".

Conclusions

In summary, this is the first study investigating social learning using an ecologically valid experimental paradigm in a large gender- and age-matched sample. We conclude that implicit and spontaneous learning from

socioemotional outcomes is not generally impaired in autism, contrary to the predictions of the social motivation theory. Nonetheless, diminished motivation or other mechanisms, such as difficulties in extracting action-outcome regularities in the environment, may be specifically associated with autism severity in females with ASC, when they do not have an explicit strategy for adapting to their social environment. This provides a strong argument for directing future research toward the clarifying gender differences and biases in ASC, as the underlying neurocognitive atypicalities in females vs. males with ASC are likely to be consistently different. In the future, this will hopefully orient clinical diagnosis and intervention to take into account individual differences in their full complexity, including the presence of co-occurring comorbidities, in particular anxiety and depression, when dealing with social atypicalities in autism.

Abbreviations

ASC	Autism spectrum condition
TD	Typically Developing
AFT	Angry feedback task
HFT	Happy Feedback Task
STAI-S-6	Short version of the State Anxiety questionnaire
STAI-T-6	Short version of the Trait Anxiety questionnaire
CATI	Comprehensive Autistic Trait Inventory
BIS-BAS	Behavioral Inhibition-Activation Scales
PHQ-8	Eight-item Patient Health Questionnaire depression scale
DSM	Diagnostic and Statistical Manual of Mental Disorders
ANOVA	Analysis of Variance
F-NE	Female participants who did not report an explicit learning strategy

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13229-024-00610-8.

Supplementary Material 1

Declarations

Author details
[1]Laboratoire sur les Interactions Cognition, Action, Émotion (LICAÉ), Université Paris Nanterre, 200 avenue de La République, Nanterre Cedex 92001, France
[2]Sorbonne Université, INSPE, Paris, France
[3]Laboratoire de Psychologie du Développement et de l'Éducation de l'enfant (LaPsyDÉ), Université Paris Cité, CNRS, 46 rue Saint-Jacques, Paris 75005, France
[4]Cognitive and Computational Neuroscience Laboratory (LNC2), Inserm U960, Department of Cognitive Studies, École Normale Supérieure, PSL University, 29 rue d'Ulm, Paris 75005, France

References
1. Lord C, Brugha TS, Charman T, Cusack J, Dumas G, Frazier T, et al. Autism spectrum disorder. Nat Rev Dis Primer. 2020;6:1–23.
2. APA APA. Diagnostic and statistical manual of mental disorders. Am Psychiatr Assoc. 2013.
3. Volkmar FR, Reichow B. Autism in DSM-5: progress and challenges. Mol Autism. 2013;4:1–6.
4. Trevisan DA, Roberts N, Lin C, Birmingham E. How do adults and teens with self-declared Autism Spectrum disorder experience eye contact? A qualitative analysis of first-hand accounts. PLoS ONE. 2017;12:e0188446.
5. Keifer CM, Mikami AY, Morris JP, Libsack EJ, Lerner MD. Prediction of Social Behavior in Autism Spectrum disorders: Explicit Versus Implicit Social Cognition. Autism Int J Res Pract. 2020;24:1758–72.
6. Lord C, Risi S, Lambrecht L, Cook EH, Leventhal BL, DiLavore PC, et al. The Autism Diagnostic Observation Schedule—Generic: a standard measure of social and communication deficits associated with the spectrum of autism. J Autism Dev Disord. 2000;30:205–23.
7. Mundy P, Sullivan L, Mastergeorge AM. A parallel and distributed-processing model of joint attention, social cognition and autism. Autism Res. 2009;2:2–21.
8. Pfeiffer D, Holingue C, Dillon E, Kalb L, Reetzke R, Landa R. Parental concerns of children with ASD by age: a qualitative analysis. Res Autism Spectr Disord. 2021;86:101817.
9. Grèzes J, Wicker B, Berthoz S, De Gelder B. A failure to grasp the affective meaning of actions in autism spectrum disorder subjects. Neuropsychologia. 2009;47:1816–25.
10. Ioannou C, Zein ME, Wyart V, Scheid I, Amsellem F, Delorme R, et al. Shared mechanism for emotion processing in adolescents with and without autism. Sci Rep. 2017;7:42696.
11. Kovarski K, Mennella R, Wong SM, Dunkley BT, Taylor MJ, Batty M. Enhanced early visual responses during Implicit Emotional faces Processing in Autism Spectrum Disorder. J Autism Dev Disord. 2019;49:871–86.
12. Kovarski K, Charpentier J, Houy-Durand E, Batty M, Gomot M. Emotional expression visual mismatch negativity in children. Dev Psychobiol. 2022;64:e22326.
13. Mennella R, Leung RC, Taylor MJ, Dunkley BT. Disconnection from others in autism is more than just a feeling: whole-brain neural synchrony in adults during implicit processing of emotional faces. Mol Autism. 2017;8:1–12.
14. Sasson NJ. The development of face processing in autism. J Autism Dev Disord. 2006;36:381–94.

15. Webb SJ, Jones EJ, Merkle K, Namkung J, Toth K, Greenson J, et al. Toddlers with elevated autism symptoms show slowed habituation to faces. Child Neuropsychol. 2010;16:255–78.

16. Yeung MK. A systematic review and meta-analysis of facial emotion recognition in autism spectrum disorder: the specificity of deficits and the role of task characteristics. Neurosci Biobehav Rev. 2022;133:104518.

17. Chevallier C, Kohls G, Troiani V, Brodkin ES, Schultz RT. The Social Motivation Theory of Autism. Trends Cogn Sci. 2012;16:231–9.

18. Dawson G, Webb SJ, McPartland J. Understanding the nature of face processing impairment in autism: insights from behavioral and electrophysiological studies. Dev Neuropsychol. 2005;27:403–24.

19. DeMayo MM, Young LJ, Hickie IB, Song YJC, Guastella AJ. Circuits for social learning: a unified model and application to Autism Spectrum Disorder. Neurosci Biobehav Rev. 2019;107:388–98.

20. Abrams DA, Padmanabhan A, Chen T, Odriozola P, Baker AE, Kochalka J, et al. Impaired voice processing in reward and salience circuits predicts social communication in children with autism. eLife. 2019;8:e39906.

21. Kohls G, Antezana L, Mosner M, Schultz R, Yerys B. Altered reward system reactivity for personalized circumscribed interests in autism. Mol Autism. 2018;9.

22. Shephard E, Milosavljevic B, Mason L, Elsabbagh M, Tye C, Gliga T, et al. Neural and behavioural indices of face processing in siblings of children with autism spectrum disorder (ASD): a longitudinal study from infancy to mid-childhood. Cortex. 2020;127:162–79.

23. Supekar K, Kochalka J, Schaer M, Wakeman H, Qin S, Padmanabhan A, et al. Deficits in mesolimbic reward pathway underlie social interaction impairments in children with autism. Brain J Neurol. 2018;141:2795–805.

24. Hedger N, Dubey I, Chakrabarti B. Social orienting and social seeking behaviors in ASD. A meta analytic investigation. Neurosci Biobehav Rev. 2020;119:376–95.

25. Jaswal VK, Akhtar N. Being versus appearing socially uninterested: challenging assumptions about social motivation in autism. Behav Brain Sci. 2019;42:e82.

26. Baumeister S, Moessnang C, Bast N, Hohmann S, Aggensteiner P, Kaiser A, et al. Processing of social and monetary rewards in autism spectrum disorders. Br J Psychiatry. 2023;222:100–11.

27. Cox A, Kohls G, Naples AJ, Mukerji CE, Coffman MC, Rutherford HJV, et al. Diminished social reward anticipation in the broad autism phenotype as revealed by event-related brain potentials. Soc Cogn Affect Neurosci. 2015;10:1357–64.

28. Keifer CM, Day TC, Hauschild KM, Lerner MD. Social and nonsocial reward anticipation in typical development and autism spectrum disorders: current status and future directions. Curr Psychiatry Rep. 2021;23:1–6.

29. Kohls G, Schulte-Rüther M, Nehrkorn B, Müller K, Fink GR, Kamp-Becker I, et al. Reward system dysfunction in autism spectrum disorders. Soc Cogn Affect Neurosci. 2013;8:565–72.

30. Matyjek M, Bayer M, Dziobek I. Autistic Traits Affect Reward Anticipation but not Reception. Sci Rep [Internet]. 2020 [cited 2023 Sep 22];10. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7242422/.

31. McPartland JC, Crowley MJ, Perszyk DR, Mukerji CE, Naples AJ, Wu J, et al. Preserved reward outcome processing in ASD as revealed by event-related potentials. J Neurodev Disord. 2012;4:16.

32. Bottini S. Social reward processing in individuals with autism spectrum disorder: a systematic review of the social motivation hypothesis. Res Autism Spectr Disord. 2018;45:9–26.

33. Clements CC, Zoltowski AR, Yankowitz LD, Yerys BE, Schultz RT, Herrington JD. Evaluation of the Social Motivation Hypothesis of Autism: a systematic review and Meta-analysis. JAMA Psychiatry. 2018;75:797–808.

34. Dichter GS. Motivational impairments in Autism May be broader Than previously thought. JAMA Psychiatry. 2018;75:773.

35. Matyjek M, Bayer M, Dziobek I. Reward responsiveness in autism and autistic traits – evidence from neuronal, autonomic, and behavioural levels. NeuroImage Clin. 2023;38:103442.

36. Schmitz N, Rubia K, van Amelsvoort T, Daly E, Smith A, Murphy DGM. Neural correlates of reward in autism. Br J Psychiatry J Ment Sci. 2008;192:19–24.

37. Xie H, Moraczewski D, McNaughton KA, Warnell KR, Alkire D, Merchant JS et al. Social reward network connectivity differs between autistic and neurotypical youth during social interaction. bioRxiv. 2023;2023.06.05.543807.

38. Jasmin K, Gotts SJ, Xu Y, Liu S, Riddell CD, Ingeholm JE, et al. Overt social interaction and resting state in young adult males with autism: core and contextual neural features. Brain. 2019;142:808–22.

39. Choi U-S, Kim S-Y, Sim HJ, Lee S-Y, Park S-Y, Jeong J-S, et al. Abnormal brain activity in social reward learning in children with autism spectrum disorder: an fMRI study. Yonsei Med J. 2015;56:705–11.

40. Kinard JL, Mosner MG, Greene RK, Addicott M, Bizzell J, Petty C, et al. Neural mechanisms of social and nonsocial reward prediction errors in adolescents with Autism Spectrum Disorder. Autism Res off J Int Soc Autism Res. 2020;13:715–28.

41. Lin A, Adolphs R, Rangel A. Impaired Learning of Social Compared to Monetary Rewards in Autism. Front Neurosci [Internet]. 2012 [cited 2023 Oct 25];6. https://www.frontiersin.org/articles/https://doi.org/10.3389/fnins.2012.00143.

42. Scott-Van Zeeland AA, Dapretto M, Ghahremani DG, Poldrack RA, Bookheimer SY. Reward processing in autism. Autism Res off J Int Soc Autism Res. 2010;3:53–67.

43. Kruppa JA, Gossen A, Oberwelland Weiß E, Kohls G, Großheinrich N, Cholemkery H, et al. Neural modulation of social reinforcement learning by intranasal oxytocin in male adults with high-functioning autism spectrum disorder: a randomized trial. Neuropsychopharmacology. 2019;44:749–56.

44. Cannon J, O'Brien AM, Bungert L, Sinha P. Prediction in Autism Spectrum disorder: a systematic review of empirical evidence. Autism Res off J Int Soc Autism Res. 2021;14:604–30.

45. Gomot M, Wicker B. A challenging, unpredictable world for people with autism spectrum disorder. Int J Psychophysiol. 2012;83:240–7.

46. Pellicano E, Burr D. When the world becomes 'too real': a bayesian explanation of autistic perception. Trends Cogn Sci. 2012;16:504–10.

47. Sapey-Triomphe L-A, Weilnhammer VA, Wagemans J. Associative learning under uncertainty in adults with autism: intact learning of the cue-outcome contingency, but slower updating of priors. Autism. 2022;26:1216–28.

48. Sinha P, Kjelgaard MM, Gandhi TK, Tsourides K, Cardinaux AL, Pantazis D, et al. Autism as a disorder of prediction. Proc Natl Acad Sci. 2014;111:15220–5.

49. Monroy C, Meyer M, Gerson S, Hunnius S. Statistical learning in social action contexts. PLoS ONE. 2017;12:e0177261.

50. Chambon V, Farrer C, Pacherie E, Jacquet PO, Leboyer M, Zalla T. Reduced sensitivity to social priors during action prediction in adults with autism spectrum disorders. Cognition. 2017;160:17–26.

51. Matyjek M, Meliss S, Dziobek I, Murayama K. A Multidimensional View on Social and Non-Social Rewards. Front Psychiatry [Internet]. 2020 [cited 2023 Oct 25];11. https://www.frontiersin.org/articles/https://doi.org/10.3389/fpsyt.2020.00818.

52. Callenmark B, Kjellin L, Rönnqvist L, Bölte S. Explicit versus implicit social cognition testing in autism spectrum disorder. Autism. 2014;18:684–93.

53. Helt MS, Fein DA, Vargas JE. Emotional contagion in children with autism spectrum disorder varies with stimulus familiarity and task instructions. Dev Psychopathol. 2020;32:383–93.

54. Barkus E, Badcock JC. A Transdiagnostic Perspective on Social Anhedonia. Front Psychiatry [Internet]. 2019 [cited 2023 Oct 25];10. https://www.frontiersin.org/articles/https://doi.org/10.3389/fpsyt.2019.00216.

55. Grèzes J, Risch N, Courtet P, Olié E, Mennella R. Depression and approach-avoidance decisions to emotional displays: the role of anhedonia. Behav Res Ther. 2023;164:104306.

56. Barkus E. The effects of Anhedonia in Social Context. Curr Behav Neurosci Rep. 2021;8:77–89.

57. Kashdan TB. Social anxiety spectrum and diminished positive experiences: theoretical synthesis and meta-analysis. Clin Psychol Rev. 2007;27:348–65.

58. Grèzes J, Erblang M, Vilarem E, Quiquempoix M, Van Beers P, Guillard M, et al. Impact of total sleep deprivation and related mood changes on approach-avoidance decisions to threat-related facial displays. Sleep. 2021;44:zsab186.

59. Hollocks MJ, Lerh JW, Magiati I, Meiser-Stedman R, Brugha TS. Anxiety and depression in adults with autism spectrum disorder: a systematic review and meta-analysis. Psychol Med. 2019;49:559–72.

60. Adams D, Ambrose K, Simpson K, Malone S, Dargue N. The relationshipbetween anxiety and social outcomes in autistic children and adolescents: a meta-analysis. Clin Child Fam Psychol Rev. 2023;26:706–20.

61. Duan S, Lee M, Wolf J, Naples AJ, McPartland JC. Higher depressive symptoms predict lower social adaptive functioning in children and adolescents with ASD. J Clin Child Adolesc Psychol off J Soc Clin Child Adolesc Psychol Am Psychol Assoc Div. 2022;53:51:203–10.

62. Uljarević M, Hedley D, Rose-Foley K, Magiati I, Cai RY, Dissanayake C, et al. Anxiety and depression from adolescence to Old Age in Autism Spectrum Disorder. J Autism Dev Disord. 2020;50:3155–65.

63. Sedgewick F, Hill V, Yates R, Pickering L, Pellicano E. Gender differences in the social motivation and friendship experiences of autistic and non-autistic adolescents. J Autism Dev Disord. 2016;46:1297–306.

64. Distefano A, Jackson F, Levinson AR, Infantolino ZP, Jarcho JM, Nelson BD. A comparison of the electrocortical response to monetary and social reward. Soc Cogn Affect Neurosci. 2018;13:247–55.

65. Lawrence KE, Hernandez LM, Eilbott J, Jack A, Aylward E, Gaab N, et al. Neural responsivity to social rewards in autistic female youth. Transl Psychiatry. 2020;10:178.

66. Spreckelmeyer KN, Krach S, Kohls G, Rademacher L, Irmak A, Konrad K, et al. Anticipation of monetary and social reward differently activates meso-limbic brain structures in men and women. Soc Cogn Affect Neurosci. 2009;4:158–65.

67. Mandy W, Charman T, Gilmour J, Skuse D. Toward specifying pervasive developmental disorder—not otherwise specified. Autism Res. 2011;4:121–31.

68. Waterhouse L. Heterogeneity thwarts autism explanatory power: a proposal for endophenotypes. Front Psychiatry. 2022;13:947653.

69. Mennella R, Bavard S, Mentec I, Grèzes J. Spontaneous instrumental avoidance learning in social contexts. Sci Rep. 2022;12:17528.

70. Kim K, Rosenthal MZ, Gwaltney M, Jarrold W, Hatt N, McIntyre N, et al. A virtual joy-stick study of emotional responses and social motivation in children with autism spectrum disorder. J Autism Dev Disord. 2015;45:3891–9.

71. Silva C, Da Fonseca D, Esteves F, Deruelle C. Motivational approach and avoidance in autism spectrum disorder: a comparison between real photographs and cartoons. Res Autism Spectr Disord. 2015;17:13–24.

72. Ho D, Imai K, King G, Stuart E, Whitworth A. Package 'MatchIt.'VersionGoogle Sch. 2018.

73. de Leeuw JR, jsPsych. A JavaScript library for creating behavioral experiments in a web browser. Behav Res Methods. 2015;47:1–12.

74. Fioravanti-Bastos ACM, Cheniaux E, Landeira-Fernandez J. Development and validation of a short-form version of the Brazilian state-trait anxiety inventory. Psicol Reflex E Crítica. 2011;24:485–94.

75. English MCW, Gignac GE, Visser TAW, Whitehouse AJO, Enns JT, Maybery MT. The Comprehensive Autistic Trait Inventory (CATI): development and validation of a new measure of autistic traits in the general population. Mol Autism. 2021;12:37.

76. Carver CS, White TL. Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the BIS/BAS scales. J Pers Soc Psychol. 1994;67:319–33.

77. Kroenke K, Strine TW, Spitzer RL, Williams JBW, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. J Affect Disord. 2009;114:163–73.

78. Mennella R, Vilarem E, Grèzes J. Rapid approach-avoidance responses to emotional displays reflect value-based decisions: neural evidence from an EEG study. NeuroImage. 2020;222:117253.

79. Vilarem E, Armony JL, Grèzes J. Action opportunities modulate attention allocation under social threat. Emotion. 2020;20:890.

80. Cuthbert BN, Insel TR. Toward the future of psychiatric diagnosis: the seven pillars of RDoC. BMC Med. 2013;11.

81. Insel T, Cuthbert B, Garvey M, Heinssen R, Pine DS, Quinn K, et al. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. Am J Psychiatry. 2010;167:748–51.

82. Mandy W. The Research Domain Criteria: A new dawn for neurodiversity research? Autism. 2018;22:642–4.

83. Hossain MM, Khan N, Sultana A, Ma P, McKyer ELJ, Ahmed HU, et al. Prevalence of comorbid psychiatric disorders among people with autism spectrum disorder: an umbrella review of systematic reviews and meta-analyses. Psychiatry Res. 2020;287:112922.

84. Khachadourian V, Mahjani B, Sandin S, Kolevzon A, Buxbaum JD, Reichenberg A, et al. Comorbidities in autism spectrum disorder and their etiologies. Transl Psychiatry. 2023;71.

85. Revelle W, Revelle MW. Package 'psych'. Compr R Arch Netw. 2015;337.

86. Lüdecke D, Ben-Shachar MS, Patil I, Makowski D. Extracting, Computing and exploring the parameters of statistical models using R. J Open Source Softw. 2020;5:2445.

87. R Core Team. R: A Language and Environment for Statistical Computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing. 2021. https://www.R-project.org/.

88. RStudio Team. RStudio: Integrated Development Environment for R [Internet]. Boston, MA: RStudio, PBC. 2020. http://www.rstudio.com/.

89. JASP Team. JASP (Version 0.18.3)[Computer software] [Internet]. 2024. https://jasp-stats.org/.

90. Brown J, Aczel B, Jiménez L, Kaufman SB, Grant KP. Intact implicit learning in autism spectrum conditions. Q J Exp Psychol. 2010;63:1789–812.

91. Foti F, De Crescenzo F, Vivanti G, Menghini D, Vicari S. Implicit learning in individuals with autism spectrum disorders: a meta-analysis. Psychol Med. 2015;45:897–910.

92. Cook J, Hull L, Crane L, Mandy W. Camouflaging in autism: a systematic review. Clin Psychol Rev. 2021;89:102080.

93. Hull L, Petrides KV, Allison C, Smith P, Baron-Cohen S, Lai M-C, et al. Putting on my best normal: social camouflaging in adults with Autism Spectrum conditions. J Autism Dev Disord. 2017;47:2519–34.

94. Lai M-C, Lombardo MV, Ruigrok AN, Chakrabarti B, Auyeung B, Szatmari P, et al. Quantifying and exploring camouflaging in men and women with autism. Autism. 2017;21:690–702.

95. Livingston LA, Happé F. Conceptualising compensation in neurodevelopmental disorders: reflections from autism spectrum disorder. Neurosci Biobehav Rev. 2017;80:729–42.

96. Hull L, Lai M-C, Baron-Cohen S, Allison C, Smith P, Petrides KV, et al. Gender differences in self-reported camouflaging in autistic and non-autistic adults. Autism Int J Res Pract. 2020;24:352–63.

97. Baron-Cohen S, Wheelwright S, Skinner R, Martin J, Clubley E. The autism-spectrum quotient (AQ): evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. J Autism Dev Disord. 2001;31:5–17.

98. Lacroix A, Nalborczyk L, Dutheil F, Kovarski K, Chokron S, Garrido M, et al. High spatial frequency filtered primes hastens happy faces categorization in autistic adults. Brain Cogn. 2021;155:105811.

99. Lacroix A, harquel sylvain, Mermillod M, Garrido M, Barbosa L, Vercueil L et al. Neural specificity of autistic women during social stimuli predictions [Internet]. PsyArXiv; 2023 Jun. https://osf.io/szqf8.

100. Goldman S, Opinion. Sex, gender and the diagnosis of autism—A biosocial view of the male preponderance. Res Autism Spectr Disord. 2013;7:675–9.

101. Krahn TM, Fenton A. The Extreme Male Brain Theory of Autism and the potential adverse effects for boys and girls with Autism. J Bioethical Inq. 2012;9:93–103.

102. Lai M-C, Lombardo MV, Pasco G, Ruigrok ANV, Wheelwright SJ, Sadek SA et al. JG Scott editor 2011 A behavioral comparison of male and female adults with high Functioning Autism Spectrum conditions. PLoS ONE 6 e20835.

103. Tillmann J, Ashwood K, Absoud M, Bölte S, Bonnet-Brilhault F, Buitelaar JK, et al. Evaluating sex and age differences in ADI-R and ADOS scores in a large European multi-site sample of individuals with Autism Spectrum Disorder. J Autism Dev Disord. 2018;48:2490–505.

104. Duvekot J, Van Der Ende J, Verhulst FC, Slappendel G, Van Daalen E, Maras A, et al. Factors influencing the probability of a diagnosis of autism spectrum disorder in girls versus boys. Autism. 2017;21:646–58.

105. McFayden TC, Putnam O, Grzadzinski R, Harrop C. Sex differences in the Developmental trajectories of Autism Spectrum Disorder. Curr Dev Disord Rep. 2023;10:80–91.

106. Hull L, Levy L, Lai M-C, Petrides K, Baron-Cohen S, Allison C, et al. Is social camouflaging associated with anxiety and depression in autistic adults? Mol Autism. 2021;12:1–13.

107. Tierney S, Burns J, Kilbey E. Looking behind the mask: Social coping strategies of girls on the autistic spectrum. Res Autism Spectr Disord. 2016;23:73–83.

108. Hull L, Mandy W. Protective effect or missed diagnosis? Females with autism spectrum disorder. Future Neurol. 2017;12:159–69.

109. Bargiela S, Steward R, Mandy W. The experiences of late-diagnosed women with autism spectrum conditions: an investigation of the female autism phenotype. J Autism Dev Disord. 2016;46:3281–94.

110. Cook A, Ogden J, Winstone N. Friendship motivations, challenges and the role of masking for girls with autism in contrasting school settings. Eur J Spec Needs Educ. 2018;33:302–15.

111. Giarelli E, Wiggins LD, Rice CE, Levy SE, Kirby RS, Pinto-Martin J, et al. Sex differences in the evaluation and diagnosis of autism spectrum disorders among children. Disabil Health J. 2010;3:107–16.

112. Guitart-Masip M, Walsh A, Dayan P, Olsson A. Anxiety associated with perceived uncontrollable stress enhances expectations of environmental volatility and impairs reward learning. Sci Rep. 2023;13:18451.

113. Vrieze E, Pizzagalli DA, Demyttenaere K, Hompes T, Sienaert P, de Boer P, et al. Reduced reward learning predicts outcome in major depressive disorder. Biol Psychiatry. 2013;73:639–45.

114. Pike AC, Robinson OJ. Reinforcement learning in patients with Mood and anxiety disorders vs Control individuals: a systematic review and Meta-analysis. JAMA Psychiatry. 2022;79:313.

115. Bishop SJ, Gagne C. Anxiety, Depression, and decision making: a computational perspective. Annu Rev Neurosci. 2018;41:371–88.

116. Loijen A, Vrijsen JN, Egger JIM, Becker ES, Rinck M. Biased approach-avoidance tendencies in psychopathology: a systematic review of their assessment and modification. Clin Psychol Rev. 2020;77:101825.

117. Radke S, Güths F, André JA, Müller BW, de Bruijn ERA. In action or inaction? Social approach–avoidance tendencies in major depression. Psychiatry Res. 2014;219:513–7.

118. Seidel E-M, Habel U, Finkelmeyer A, Schneider F, Gur RC, Derntl B. Implicit and explicit behavioral tendencies in male and female depression. Psychiatry Res. 2010;177:124–30.

119. Leung FYN, Sin J, Dawson C, Ong JH, Zhao C, Veić A, et al. Emotion recognition across visual and auditory modalities in autism spectrum disorder: a systematic review and meta-analysis. Dev Rev. 2022;63:101000.

120. Kang E, Keifer CM, Levy EJ, Foss-Feig JH, McPartland JC, Lerner MD. Atypicality of the N170 event-related potential in Autism Spectrum disorder: a Meta-analysis. Biol Psychiatry Cogn Neurosci Neuroimaging. 2018;3:657–66.

121. Pellicano E, Adams D, Crane L, Hollingue C, Allen C, Almendinger K et al. A possible threat to data integrity for online qualitative autism research. Autism. 2023;13623613231174543.

122. Lehnhardt F-G, Falter CM, Gawronski A, Pfeiffer K, Tepest R, Franklin J, et al. Sex-related cognitive profile in autism spectrum disorders diagnosed late in life: implications for the female autistic phenotype. J Autism Dev Disord. 2016;46:139–54.

123. Virués-Ortega J. Applied behavior analytic intervention for autism in early childhood: Meta-analysis, meta-regression and dose–response meta-analysis of multiple outcomes. Clin Psychol Rev. 2010;30:387–99.

124. Belcher HL, Morein-Zamir S, Mandy W, Ford RM. Camouflaging intent, First impressions, and age of ASC Diagnosis in autistic men and women. J Autism Dev Disord. 2022;52:3413–26.

125. Kohls G, Chevallier C, Troiani V, Schultz RT. Social wanting dysfunction in autism: neurobiological underpinnings and treatment implications. J Neurodev Disord. 2012;4:10.

126. Dichter GS, Felder JN, Green SR, Rittenberg AM, Sasson NJ, Bodfish JW. Reward circuitry function in autism spectrum disorders. Soc Cogn Affect Neurosci. 2012;7:160–72.

127. Müller R-A, Fishman I. Brain Connectivity and Neuroimaging of Social Networks in Autism. Trends Cogn Sci. 2018;22:1103–16.

Publisher's Note