Contents lists available at ScienceDirect

International Journal of Clinical and Health Psychology

Reduced stereotypicality and spared use of facial expression predictions for social evaluation in autism

Marta Robles [a,b], Irene Ramos-Grille [a,c], Amaia Hervás [d,e], Enric Duran-Tauleria [e],
Jordi Galiano-Landeira [a], Jolie B. Wormwood [f], Christine M. Falter-Wagner [b],
Lorena Chanes [a,g,h,*]

[a] *Department of Clinical and Health Psychology, Universitat Autònoma de Barcelona, Barcelona, Spain*
[b] *Department of Psychiatry and Psychotherapy, LMU University Hospital, LMU Munich, Germany*
[c] *Division of Mental Health, Consorci Sanitari de Terrassa, Terrassa, Catalunya, Spain*
[d] *Child and Adolescent Mental Health Service, Hospital Universitari Mútua de Terrassa, Barcelona, Spain*
[e] *Institut Global d'Atenció Integral del Neurodesenvolupament (IGAIN), Barcelona, Spain*
[f] *University of New Hampshire, Durham, NH, USA*
[g] *Institut de Neurociències, Universitat Autònoma de Barcelona, Barcelona, Spain*
[h] *Serra Húnter Programme, Generalitat de Catalunya, Barcelona, Spain*

ARTICLE INFO

ABSTRACT

Background/Objective: Autism has been investigated through traditional emotion recognition paradigms, merely investigating accuracy, thereby constraining how potential differences across autistic and control individuals may be observed, identified, and described. Moreover, the *use* of emotional facial expression information for social functioning in autism is of relevance to provide a deeper understanding of the condition.
Method: Adult autistic individuals ($n = 34$) and adult control individuals ($n = 34$) were assessed with a social perception behavioral paradigm exploring facial expression predictions and their impact on social evaluation.
Results: Autistic individuals held less stereotypical predictions than controls. Importantly, despite such differences in predictions, the *use* of such predictions for social evaluation did not differ significantly between groups, as autistic individuals relied on their predictions to evaluate others to the same extent as controls.
Conclusions: These results help to understand how autistic individuals perceive social stimuli and evaluate others, revealing a deviation from stereotypicality beyond which social evaluation strategies may be intact.

Introduction

Autism has been traditionally described as a neurodevelopmental disorder characterized by symptoms in communication and social interaction, as well as the presence of restricted and repetitive patterns of behavior, interests, or activities (Autism Spectrum Disorder; American Psychiatric Association, APA, 2022). Among other aspects related to emotion and social cognition, facial expression perception in autism has been extensively assessed. Studies have reported primarily deficits, difficulties, or poorer performance on facial expression recognition among autistic individuals (for meta-analyses see Lozier et al., 2014; Uljarevic & Hamilton, 2013; for a recent review and meta-analysis, see Yeung, 2022). Such observations, described in terms of deficits or

difficulties, are tightly linked to assumptions about "accuracy" in emotion perception; studies typically assess an individuals' ability to identify stereotypical or prototypical expressions of emotion (e.g., matching a posed, often exaggerated facial expression to one of a limited set of provided emotion words). Importantly, although autistic individuals tend to perform more poorly on such tasks than non-autistic individuals, it is still unclear *why* these differences emerge or what such measures can actually reveal about emotion perception ability in the real world at all.

Much research on emotion perception, including but not limited to research on autism, relies on assumptions about emotion stemming from what is sometimes referred to as the classical view of emotion (for an overview, see Tracy & Randles, 2011). From this perspective, a core

facial configuration ("expression") exists for each of a series of emotions that are considered biologically wired and universally shared (for a review, see, e.g., Barrett et al., 2019). Under this view, sadness is always expressed by exhibiting a pouting facial expression, happiness by a smiling facial expression, and so on. Thus, a person's ability to match a posed, stereotypical expression (i.e., a facial expression that is normatively associated with a given emotion category in a particular culture) to a given emotion label may be considered a measure of emotion perception accuracy from this perspective, because it is presumed that there is an objectively correct answer and that the process is consistent with how emotion perception unfolds in everyday life.

However, there is now substantial empirical data challenging these assumptions, as evidence shows that emotions are experienced and expressed in highly variable ways, and that they vary significantly across cultures and situations (Barrett et al., 2019). In particular, facial expressions of emotion are highly variable (see Fernández-Dols & Crivelli, 2013) and individuals fail to consistently make the stereotypical expressions used in emotion perception research at times when they are actually experiencing those emotions (Durán & Fernández-Dols, 2021). These findings are consistent with constructionist views of emotion, such as the Theory of Constructed Emotion (Barret, 2017), which posit that emotions are not biologically discrete but instead represent categories constructed by applying conceptual knowledge about emotion (learned through one's culture, language, and experience) to make meaning of ongoing sensory information arising from the body (e.g., changes in heart rate or breathing). From this perspective, emotions represent heterogeneous categories of instances, such that there is both variability within emotion categories (i.e., not all instances of anger are identical, even for the same person) and similarity across categories (i.e., instances of anger and happiness might both involve smiling or an elevated heart rate). Thus, according to constructionist theories of emotion, variability in the components of emotion, including facial expressions, are expected, undermining the assumption made in most emotion perception research that there is a one-to-one correspondence between a specific facial configuration and the experience of a given emotion (Barrett et al., 2019). Simply put, people do not always pout when sad or smile when happy. As such, typical emotion perception tasks that ask individuals to match posed facial expressions to a given emotion word cannot assess 'accuracy' in emotion perception. Instead, they measure the person's ability to identify normative or stereotypical expressions that belong to a given emotion category in a given culture. Here, we move away from assessing 'accuracy' in emotion perception to examine the extent to which autistic individuals expect to see stereotypical facial expressions in different emotional evocative scenarios.

Indeed, amongst constructivist views, expectations or predictions are posited to be the basis of emotion perception according to The Theory of Constructed Emotion (Barret, 2017), which builds on predictive processing accounts of perception and action, suggesting that emotion experience and perception unfold predictively, with prior experience and beliefs guiding experience and perception in critical ways (Barrett, 2017; Hoemann et al., 2020). In recent years, multiple brain-related conditions, including autism, have been described in terms of predictive processing or predictive coding (for a recent review, see, Smith et al., 2021). The core idea of predictive processing is that the brain constantly issues predictions about what will happen next based on previous experiences, which are then compared to actual sensory input from the body and the external world (see, e.g., Friston, 2005, 2010). When the difference between predictions and actual sensory input, called prediction error, is minimal, *predictions* are thought to drive perception and behavior. Only when the prediction error passes some threshold does actual sensory input more strongly inform perception and behavior, and the model used to issue predictions is updated to reduce such prediction error in the future. Moreover, reliance on predictions and prediction errors are weighted, i.e., are prioritized or not, based on their perceived reliability (i.e., precision; Yon & Frith, 2021). With regards to autism, several models have been proposed, roughly

suggesting a privileged processing of prediction error compared to predictions (Friston et al., 2013; Lawson et al., 2014; Pellicano & Burr, 2012; Quattrocki & Friston, 2014a; Sinha et al., 2014; van de Cruys et al., 2014; for a recent review, see, Cannon et al., 2021), leading to a model that may become too complex too soon, including details that may remain overlooked by non-autistic individuals. Practically speaking, prioritizing prediction error could lead autistic individuals to have difficulties generalizing and adapting to the constantly changing world we live in, requiring instead anticipation, sameness, and routines, a core feature of autism (American Psychiatric Association, 2022). In the context of emotion perception, an overreliance on prediction error would be consistent with expectations that are more nuanced and less normative.

In the present study, we use a social perception task in which facial expression predictions are induced through short written scenarios and subsequently confirmed or violated to assess how much prediction error a person experiences and how normative or stereotypical their predictions about facial expressions of emotion are. In the task, participants are presented with a picture of a person and read a text that describes a situation (scenario) thought to be normatively associated with a specific emotion (e.g., attending a funeral and sadness; enjoying time with friends and happiness). We ask them to imagine how the person would look in that scenario, and then show them the person's face depicting a stereotypical emotion expression that either matches or mismatches the emotion normatively associated with the scenario. Subsequently, we assess how much the face matched their prediction about what the person would look like and how much they liked the person.

In previous studies using this task, healthy adult participants were shown to hold fairly stereotypical predictions about facial expressions in emotion contexts (Chanes et al., 2018): participants rated stereotypical facial expressions matching the emotion evoked by the scenario (e.g., a smiling face for a happy scenario) as more similar to what they had imagined than when the stereotypical facial expression did not match the emotion evoked by the preceding scenario (e.g., a smiling face for a sad scenario). Critically, faces posing stereotypical expressions that were rated as more predictable (i.e., as more in line with the perceiver's expectations) were also judged as more likable and trustworthy. Thus, in emotion contexts, facial expression stereotypes seem to drive individuals' predictions about what expression they will encounter in a given context, and people are judged more favorably when their expression matches a perceiver's predictions about how they should look (i.e., the perceiver's internal model).

Using this same task, we here explore facial expression perception in adult autistic individuals, assessing the extent to which they predict to encounter stereotypical facial expressions in emotionally evocative contexts. By assessing the impact of such expectations on likability ratings, we aimed to go one step further and study whether these predictions impact social evaluation in autism, which may contribute to better understand social function in this condition. We hypothesized that predictions about facial expressions would be less stereotypical for autistic than control individuals, and that such differences could be associated with differences in social evaluation.

Materials and methods

Participants

Thirty-four autistic adult individuals (sixteen females, eighteen males; autistic group) and 34 adult individuals without a current or history of a diagnosis of any psychiatric or neurological condition and not taking psychoactive medication (sixteen females, eighteen males; control group) participated in the study. Sample size was based on previous studies with overall-healthy individuals (Chanes et al., 2018) and individuals with depression (Ramos-Grille et al., 2022), and in line with studies investigating facial expression perception in autism (see, e.g Yeung, 2022). Autistic individuals were recruited from the Global

Institute of Neurodevelopment Integrated Care (IGAIN), a healthcare center specialized in autism located in Barcelona (Spain). The diagnosis of autism spectrum disorder (ASD), as well as the absence of an intellectual developmental disorder, was confirmed by clinical experts at the center according to DSM-5 (APA, 2013) through extensive clinical evaluation. Inclusion criteria for the autistic group were: (i) a confirmed diagnosis of ASD, (ii) no history of brain injury or other neurological conditions, (iii) no presence of an intellectual developmental disorder, (iv) age $\geq$ 18 years old, (v) normal or corrected-to-normal vision, and (vi) native Spanish speaker or bilingual Catalan-Spanish. In this group, twenty-one participants had one or more comorbid diagnoses. More specifically, 13 of them presented one comorbid diagnosis and 8 presented two or more comorbid diagnoses, roughly reflecting the ratios of comorbid diagnoses in autism observed in previous literature (APA, 2022) (see Table 1 for more details). Twenty-one participants were taking one or more medications.

Control individuals were recruited in Barcelona area through word-of-mouth and advertisements shared on social media. They were selected to match autistic participants' gender and age within a range of $\pm$ 5 years. For control individuals, the same inclusion criteria applied, as well as not having been diagnosed with any psychiatric or neurological condition and not be taking psychoactive medication. One control participant was excluded after data collection but before any data analysis, because of a diagnosed psychiatric condition reported at the end of the session. This participant was replaced by another participant of the same gender and similar age before any data analysis.

The study was approved by the Autonomous University of Barcelona's Institutional Review Board. All participants gave written informed consent and did not receive any monetary compensation for their participation. There was no community involvement in the reported study.

The autistic and control groups were 28 $\pm$ 2 years old (*Mean $\pm$ SE*) and 31 $\pm$ 1 years old (*Mean $\pm$ SE*), respectively, the difference being small although significant (Mann-Whitney two-tailed *U*-test, $U = 754.5$, $p = 0.031$, *rB*$= 0.305$, 95 % *CI* [0.039, 0.531]) (Table 1).

We performed an additional analysis on control participants focusing on autistic traits. For that specific analysis we included not only control participants but also 7 control participants from a different ongoing study (with the same inclusion criteria and study characteristics), adding up to a total of 41 individuals (Autism-like traits group; see Table 1 for more details). This group was 30 $\pm$ 1 years old (*Mean $\pm$ SE*).

Experimental procedure

Data collection took place at the healthcare center (IGAIN). First, sociodemographic data was collected (reported age, sex, gender, handedness, and education level). Next, participants performed a predictive processing and social perception task (Chanes et al., 2018) after which they completed the abridged Version of the Autism-Spectrum Quotient, (AQ-Short; Hoekstra et al., 2011; Spanish version used: Lugo-Marín et al., 2019) among other questionnaires not used for this study (Positive and Negative Affect Scale, Watson et al., 1988; Emotion Regulation Questionnaire, Gross & John, 2003; short version of the UPPS-P impulsive behavior scale, Billieux et al., 2012). Diagnoses (including co-occurring conditions), and medications at the time of study participation for the autistic group were collected from medical records.

Predictive processing and social perception task

Stimuli were presented using E-Prime 3 (Psychology Software Tools, Inc., Pittsburgh, PA, USA) running on an HP ProBook 640 G4 (display size: 14'', resolution: 1920 $\times$ 1080).

Table 1

Sociodemographic and clinical variables for the autistic (*n* = 34) and control (*n* = 34) groups, as well as for the autism-like traits group (*n* = 41).

| Variables	Autistic group		Control group		Autism-like traits group	
	n/mean	%/[SD]	n/mean	%/[SD]	n/mean	%/[SD]
GENDER						
Women	16	53 %	16	53 %	19	46 %
Men	18	47 %	18	47 %	22	54 %
AGE (years)	28	[10]	31	[7]	30	[7]
EDUCATION LEVEL						
Primary or lower	1	3 %	0	0 %	0	0 %
High school	28	82 %	13	38 %	14	34 %
University	5	15 %	21	62 %	27	66 %
AQ-SHORT SCORE	72	[11]	53	[10]	54	[11]
COMORBID DIAGNOSES	21	62 %				
One comorbid diagnosis	13	38 %				
Two or more comorbid diagnoses	8	24 %				
Anxiety Disorders	17	50 %				
Generalized Anxiety Disorder	8	24 %				
Social Anxiety	9	26 %				
Depressive Disorders	4	12 %				
Major Depression	2	6 %				
Unspecified Depressive Disorder	2	6 %				
Neurodevelopmental Disorders	6	18 %				
Attention Deficit Hyperactivity Disorder	5	15 %				
Specific Learning Disorder (Dyscalculia)	1	3 %				
Other disorders						
Unspecified Bipolar Disorder	1	3 %				
Obsessive-Compulsive Disorder	2	6 %				
Post-Traumatic Stress Disorder	1	3 %				
PSYCHOLOGICAL TREATMENT	34	100 %				
PHARMACOLOGICAL TREATMENT	21	62 %				
Alpha-2 adrenergic agonist	7	21 %				
Benzodiazepines	3	9 %				
Central nervous system stimulant	1	3 %				
Norepinephrine-dopamine reuptake inhibitor	13	38 %				
Selective serotonin reuptake inhibitor	7	21 %				
Serotonin-norepinephrine reuptake inhibitors	6	18 %				
Tetracyclic antidepressants	2	6 %				
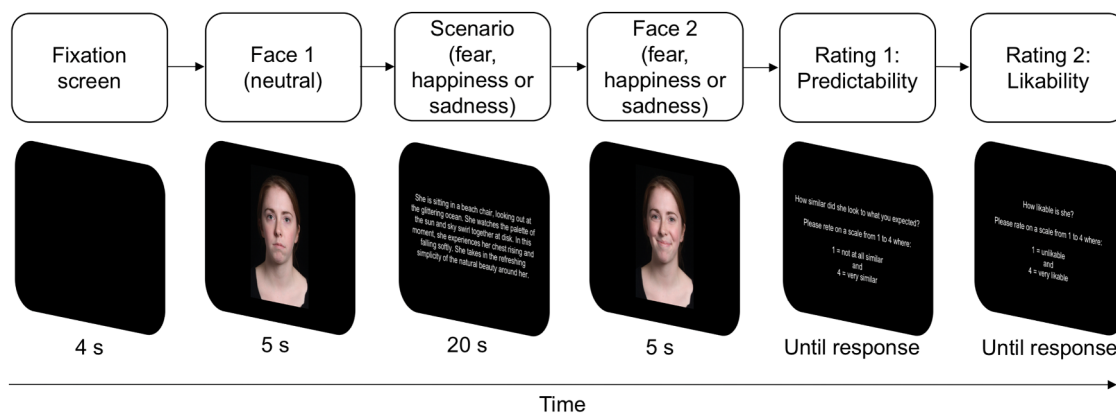
Fig. 1. Sequence of screens on a sample experimental trial. After a fixation screen, a target person displaying a neutral facial expression was presented. Next, a scenario (short text) was displayed evoking either fear, happiness, or sadness. Then, the same target person was shown exhibiting a stereotypical facial expression for fear, happiness, or sadness. The facial expression could match the stereotypical expression for the emotion evoked by the scenario (matched trials) or not (non-matched trials). Afterwards, two ratings (predictability and likability) were performed.

Each trial began with a black fixation screen (4 s) (Fig. 1). A photograph of a target person with a neutral facial expression was then shown at the center of the screen (Face 1; 5 s). Next, a short story (Scenario; 20 s) was displayed in white font. Each scenario, describing a situation experienced by the target person, was aimed to evoke one of the three following emotions: fear, happiness, or sadness. Participants were asked to imagine, while reading through the story, how the target person would look in that scenario. After the scenario, a second photograph of the target person was displayed, exhibiting a stereotypical facial expression for one of the three possible evoked emotions (fear, happiness, or sadness; Face 2; 5 s). On some trials, the stereotypical expression corresponded to the emotion evoked by the scenario (matched trials; e.g., a pouting facial expression following a scenario evoking sadness) and on some trials it did not (nonmatched trials; e.g., a pouting facial expression following a scenario evoking happiness or fear). Finally, participants were asked to complete two ratings on four-point scales with no time limit. In the first rating, they were asked to indicate how similar the target person looked (Face 2) compared to what they had imagined while reading the scenario, on a scale from 1 = *"not at all similar"* to 4 = *"very similar"* (predictability rating; Rating 1). In the second rating, participants indicated how likable the target person was, on a scale from 1 = *"very unlikable"* to 4 = *"very likable"* (likability rating; Rating 2). Thus, the task was designed to tackle subjective predictions about facial configurations. Emotion categories were not explicitly mentioned to participants at any time. We aimed to assess how well the images presented matched participants' predictions, which may be considered as a rough behavioral correlate of prediction error. Because there were no right/wrong answers, no accuracy was computed, and results may not be described in terms of a better/worse performance.

Prior to the start of the task, participants were asked to read task instructions displayed on the screen, which included sample photographs and sample rating screens. After going through the instructions and asking the researcher any questions they might had, participants were asked to verbally summarize the task to the researcher. Once the researcher confirmed that the task had been correctly understood by participants, they performed 3 practice trials. These trials were all matched trials (one evoking each of the three possible emotions: fear, happiness, and sadness). The researcher observed participants' responses to these trials. If the predictability rating on any of these three trials was relatively low (1 or 2), the researcher asked questions to the participant to confirm that the low ratings on those trials were not due to a lack of understanding of the instructions, a lack of understanding of the texts, or any other aspect suggesting a lack of capability to perform the task. The script followed by the researcher was: "I have seen that you

have responded 1 or 2 on some/all of the trials. How did you imagine/ expect that the person would look?". By the participant's response, the researcher would confirm that the low rating was due to differences in their predictions with regards to the displayed photograph, rather than a lack of understanding of the instructions, the texts or other aspects preventing the participant from performing the task adequately.

The task included 45 experimental trials divided into 5 blocks. Each block consisted of 9 trials and participants were offered to take a short break in between if they wished. The 45 experimental trials included 15 scenarios evoking each of the three emotion categories: fear, happiness, and sadness. Among the 15 scenarios evoking a given emotion, 9 scenarios were followed by the corresponding stereotypical facial expression (matched trials). Six scenarios were followed by a stereotypical facial expression for a different emotion than the one evoked by the scenario, 3 for each of the other two emotion categories (nonmatched trials). This added up to a total of 27 matched trials and 18 nonmatched trials per participant. Identities of the models shown for each scenario, scenario order, and match condition assigned to a given scenario were all pseudorandomized within participants.

A different target person (identity) was used for each of the 48 trials (3 practice trials: 2 females, 1 male; 45 experimental trials: 28 females, 17 males). Color photographs (400 × 600 pixels) of human faces with closed mouths and a direct gaze, which belong to the Interdisciplinary Affective Science Laboratory[1] (www.affective-science.org), were used (for more details, see Chanes et al., 2018). We used the scenarios used in previous studies (Chanes et al., 2018, originally developed by Wilson-- Mendenhall et al., 2013), translated into Spanish (used in Draganov et al., 2023; Ramos-Grille et al., 2022). The scenarios narrated a scene matching the target person's gender, and evoked either happiness, sadness, or fear, with high or low arousal for each emotion (e.g., "She is sitting in a beach chair, looking out at the glittering ocean. She watches the palette of the sun and sky swirl together at dusk. In this moment, she experiences her chest rising and falling softly. She takes in the refreshing simplicity of the natural beauty around her".).

Questionnaire

We used the Spanish abridged version of the Autism-Spectrum Quotient, (AQ-Short; Hoekstra et al., 2011; Spanish version used: Lugo-Marín et al., 2019) in order to measure the presence of autistic

---

traits across participants. The AQ-Short consists of 28 self-reported items describing typical autistic traits such as "I find it difficult to make new friends", with responses on a 4-point scale, with possible answers being "1 = definitely agree", "2 = slightly agree", "3 = slightly disagree", and "4 = definitely disagree". Scoring is reversed for items in which an "agree" response is not characteristic of autism and item scores are summed. Lugo-Marín et al. (2019) propose a cut-off score of 65 to identify potentially autistic individuals with a sensitivity of 0.98 and a specificity of 0.84. The results of the reliability analysis suggest a good internal structure (intraclass correlation coefficients ranging from 0.90 to 0.97), in line with the results found for the original version (Hoekstra et al., 2011).

Data analysis

Non-aggregated data from individual trials were analyzed using hierarchical linear modeling (HLM 7.0; Scientific Software International, Inc., Skokie, IL, USA). HLM analysis was used to avoid aggregation across trials and model variability in trial-by-trial performance nested within each participant. We used a continuous sampling model with participants treated as a random factor, and a restricted maximum likelihood method of estimation for model parameters (Raudenbush & Bryk, 2002). Continuous trial-level variables (e.g., predictability ratings) were centered around each participant's mean when entered as predictors in the models. Dummy-coded variables (e.g., match condition) were uncentered (Enders & Tofighi, 2007). All HLM models had random intercepts. Additionally, when analyzing traits rather than autism as a category, a linear regression model was computed. The linear regression and additional statistical analyses were carried out using JASP (JASP Team, 2022; version 0.9.2). For comparisons across control and autistic groups, we used t-tests or Mann-Whitney U-tests when normality assumptions were not met. Prism 9 (GraphPad Software, San Diego, CA, USA) was used for data visualization.

Results

Self-reported autistic traits for autistic and control groups

As expected, AQ-Short scores differed significantly across groups (Mann-Whitney two-tailed U-test, $U = 110$, $p < 0.001$, $rB = -0.840$, 95 % CI [0.883, -0.681]), the autistic group presenting more autistic traits (*Mean ± SE*: 72 ± 2) than the control group (*Mean ± SE*: 53 ± 2) (Table 1). AQ-Short scores from one autistic individual were missing due to tiredness and inability to finalize responses to the questionnaire. The autism-like traits group had AQ-Short scores of 54 ± 2 (*Mean ± SE*).

Predictability of facial expressions of emotion

Overall

An HLM analysis was computed, with trial-level predictability ratings as the outcome variable and group (autistic and control) as a dummy-coded subject-level predictor variable. This analysis revealed that, overall, autistic individuals rated facial expressions as significantly less similar to what they expected (lower predictability ratings) (*Mean ± SE*: 2.49 ± 0.07) than control individuals (Mean ± SE: 2.65 ± 0.04) ($B = -0.17$, $SE = 0.08$, $t(66) = 2.16$, $p = 0.034$). This is consistent with autistic individuals having increased prediction error relative to controls.

By match condition

Next, we conducted an HLM analysis with trial-level predictability ratings as the outcome variable, match condition as a dummy-coded trial-level predictor variable, and group (autistic and control) as a dummy-coded subject-level predictor variable, with the two-way interaction term included. As expected, matched facial expressions (stereotypical facial expressions for the evoked emotion) were rated as significantly more predictable (*Mean ± SE*: 3.25 ± 0.05) than

nonmatched facial expressions (*Mean ± SE*: 1.76 ± 0.05) for the control group ($B = 1.49$, $SE = 0.06$, $t(66) = 24.85$, $p < 0.001$), which is consistent with prior studies (Chanes et al., 2018; Draganov et al., 2023; Ramos-Grille et al., 2022). Autistic individuals also rated matched facial expressions as significantly more predictable (*Mean ± SE*: 2.93 ± 0.09) than nonmatched (*Mean ± SE*: 1.82 ± 0.08) facial expressions ($B = 1.11$, $SE = 0.10$, $t(66) = 11.05$, $p < 0.001$). However, a significant interaction indicated that the difference in predictability ratings for matched vs. nonmatched facial expressions (namely 'match effect'), was significantly less pronounced among autistic individuals than among control individuals ($B = -0.39$, $SE = 0.12$, $t(66) = 3.30$, $p = 0.002$).

To explore this interaction further, we separately analyzed predictability ratings for trials with matched and nonmatched facial expressions. These analyses revealed that autistic individuals rated the expressions as significantly less predictable than control individuals on trials where the facial expression matched the stereotypical expression for the evoked emotion (matched trials; $B = 0.32$, $SE = 0.10$, $t(66) = 3.25$, $p = 0.002$), but there were no differences in predictability ratings on trials where the facial expression did not match the stereotypical expression for the evoked emotion (nonmatched trials; $B = -0.06$, $SE = 0.09$, $t(66) = 0.69$, $p = 0.490$). This pattern of results suggests that autistic individuals have less stereotypical predictions for facial expressions of emotion than those in the control group; they found stereotypical expressions matched to the emotion evoked by the scenario (i. e., smiling faces following a normatively happy scenario) less in line with their own predictions for what the target person would look like in that scenario than did controls.

By match and emotion condition

We further explored whether the observed differences between autistic individuals vs. control individuals in terms of predictability ratings were consistent across the three different emotion categories evoked by the scenarios (see Fig. 2). To do so, we conducted the above analyses separately for trials with each of the three different evoked emotions (happiness, sadness, and fear). These analyses revealed that, compared to controls, autistic individuals exhibited a less pronounced difference in predictability ratings on matched vs. nonmatched trials for scenarios evoking happiness (autistic individuals (*Mean ± SE*): 3.28 ± 0.09 vs. 1.53 ± 0.08; control individuals (*Mean ± SE*): 3.54 ± 0.05 vs. 1.36 ± 0.07; $B = -0.43$, $SE = 0.15$, $t(66) = 2.87$, $p = 0.006$), sadness (autistic individuals (*Mean ± SE*): 2.75 ± 0.10 vs. 1.97 ± 0.09; control individuals (*Mean ± SE*): 3.18 ± 0.07 vs. 1.93 ± 0.07; $B = -0.48$, $SE = 0.13$, $t(66) = 3.71$, $p < 0.001$), and fear (autistic individuals (*Mean ± SE*): 2.76 ± 0.10 vs. 1.97 ± 0.10; control individuals (*Mean ± SE*): 3.03 ± 0.07 vs. 1.99 ± 0.07; $B = -0.25$, $SE = 0.15$, $t(66) = 1.61$, $p = 0.112$), though this interaction only reached significance for scenarios evoking happiness and sadness. Further, analyses revealed that autistic individuals rated matched expressions as significantly less predictable than controls for all three emotion scenarios (happiness: $B = 0.26$, $SE = 0.10$, $t(66) = 2.46$, $p = 0.016$; sadness: $B = 0.44$, $SE = 0.13$, $t(66) = 3.49$, $p < 0.001$; fear: $B = 0.27$, $SE = 0.12$, $t(66) = 2.25$, $p = 0.028$). No differences in predictability ratings between autistic individuals and control participants were observed for nonmatched trials for any emotion condition (happiness: $B = -0.17$, $SE = 0.10$, $t(66) = 1.64$, $p = 0.105$; sadness: $B = -0.04$, $SE = 0.11$, $t(66) = 0.36$, $p = 0.722$; fear: $B = 0.02$, $SE = 0.12$, $t(66) = 0.17$, $p = 0.869$). Taken together, these results indicate that the effects observed do not seem to depend on the specific evoked emotion but rather emerge across emotion categories.

Facial expression predictions with regards to autism-like traits

We explored if the above-described reduced match effect (i.e., difference in predictability ratings for matched vs. nonmatched facial expressions) for autistic vs. control individuals is also observed when considering autism-like traits in non-autistic individuals. Only for this additional analysis, data from our control sample ($N = 34$) as well as
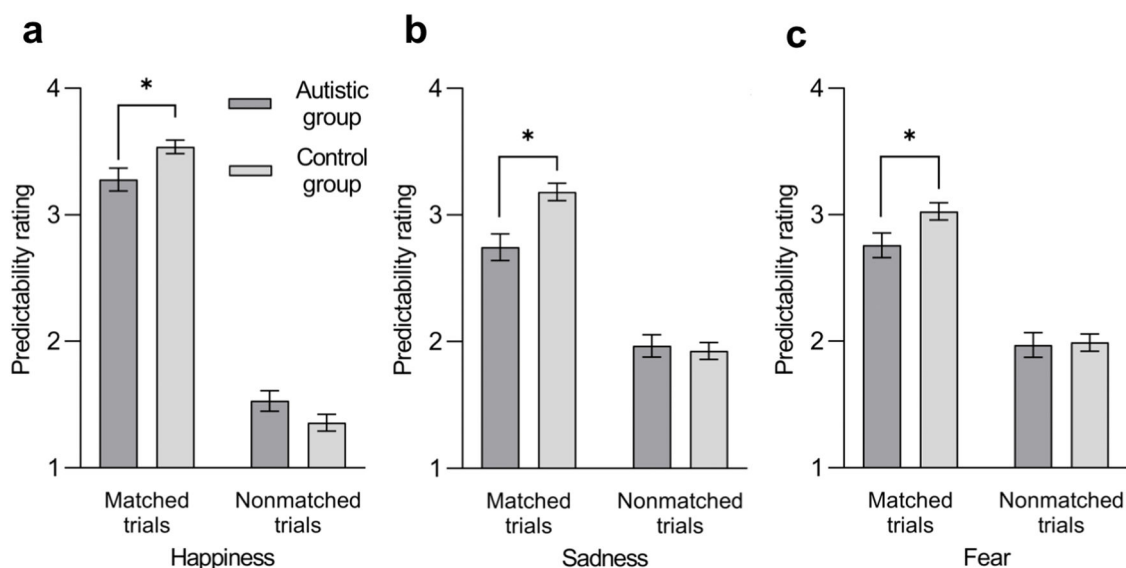
Fig. 2. Predictability ratings for autistic and control individuals by scenario evoked emotion and match condition. Means and standard errors for the autistic (dark gray) and the control (light gray) groups. Match effect (difference between ratings on matched and nonmatched trials) was reduced in autistic vs. control individuals. This effect was driven primarily by differences in predictability ratings on matched trials. Asterisks indicate $p < 0.05$.

from 7 further control participants from a different ongoing study were used (total $N = 41$). A match effect variable was computed for each participant by subtracting the mean of each participant's predictability ratings on matched trials and nonmatched trials. We then performed a linear regression with AQ-Short scores as the predictor variable and match effect as the outcome variable. This analysis revealed that, for non-autistic individuals, higher autism-like traits were associated with a smaller match effect ($R^2 = 0.109$; $F(1, 39) = 4.767$, $t = -2.183$ p = 0.035). Thus, the association between autism and having less stereotypical predictions is observed both when comparing those with and without an autism diagnosis, as well as for non-autistic individuals with more (vs. fewer) autism-like traits.

Predictability of facial expressions and social evaluation

The impact of facial expression predictability on social evaluation was assessed as the relationship between predictability and likability ratings using an HLM analysis with trial-level likability ratings as the outcome variable, predictability ratings as a trial-level predictor variable, and group (autistic and control individuals) as a dummy-coded subject-level predictor variable, with the two-way interaction term included in the model. Consistent with prior work (Chanes et al., 2018), this analysis revealed a positive relationship between predictability and likability ratings across conditions. Facial expressions that were rated as more predictable were also rated as more likable for both groups (control individuals: $B = 0.23$, $SE = 0.03$, $t(66) = 6.70$, $p < 0.001$; autistic individuals: $B = 0.25$, $SE = 0.04$, $t(66) = 6.50$, $p < 0.001$). Importantly, this relationship did not differ significantly between groups ($B = -0.01$, $SE = 0.05$, $t(66) = 0.25$, $p = 0.801$). This pattern of results suggests that individuals like others more when they display predicted expressions (i. e., expressions that better match the perceiver's own prediction of what someone will look like in a given scenario), and this relationship between predictability and liking does not differ across autistic individuals and control individuals.

Discussion

Compared to controls, autistic individuals reported greater inconsistencies between stereotypical facial expressions of emotion and their own predictions about what someone's facial expression would be in a given context, suggesting their predictions are less stereotypical

than controls'. Similarly to controls, however, autistic individuals liked others more when their facial expression better matched their own predictions about what it would look like in a given scenario.

Although the autistic group had lower predictability ratings than the control group overall, i.e., a larger perceived difference between what they expected and the actual stimuli displayed, this effect was driven by ratings in the matched condition only; rather than being systematically lower across all trials, predictability ratings for autistic individuals were lower specifically for matched trials, where the stereotypical facial expression for the emotion depicted by the scenario was displayed. There were no differences observed between the autistic and control groups in terms of predictability ratings, for nonmatched trials, where a stereotypical facial expression for a different emotion than that depicted by the scenario was displayed. That we observed similar predictability ratings for nonmatched trials across groups indicates that both groups found categorically normative mismatches (i.e., a smiling face in a sad scenario) unexpected to the same extent. Thus, findings point to autistic individuals having less stereotypical predictions when thinking about what someone's facial expression will be in a given emotionally evocative situation. An alternative possibility is that autistic individuals have a general "deficit", i.e., did not have clear facial expression predictions in emotion contexts, though this explanation seems less likely given our rigorous protocol for ensuring understanding and compliance with task instructions and ability to perform it among all participants. Critically, autistic individuals also had significantly higher predictability ratings in matched than nonmatched trials (although this difference was smaller among autistic participants than controls), suggesting they were making predictions to some extent as otherwise predictability ratings would be expected not to differ significantly across both types of trials within the autistic group. Moreover, we found converging evidence looking at autism-like traits (as assessed by the AQ-Short) only in non-autistic individuals: individuals that showed higher characteristics of autism also showed a smaller difference in predictability ratings between matched and nonmatched trials. Furthermore, autistic individuals relied on their predictions to the same extent than controls to rate likability. Nevertheless, we did not directly assess the content of participants' predictions nor their confidence in them, so future work should examine the extent to which these qualities of the predictions differ across autistic and control participants in general as well as in stereotypical contexts specifically.

This lack of stereotypical predictions for facial expressions of

emotion among autistic individuals is consistent with observations of lower "accuracy" for autistic vs. control individuals in previous studies using traditional emotion recognition tasks framed in classical views of emotion perception (for meta-analyses see Lozier et al., 2014; Uljarevic & Hamilton, 2013; Yeung, 2022). In these tasks, the facial expression stipulated as correct for a given emotion category is typically a highly stereotypical expression for that emotion. To the extent that autistic individuals do not expect expressions as stereotypical as non-autistic individuals, this could lead to less accurate matching of expressions and emotion words in a typical emotion perception task.

Our findings are also consistent with recent predictive coding accounts of autism that suggest a privileged processing of prediction error in autistic individuals (for a recent review, see Cannon et al., 2021). This overreliance on prediction error may lead autistic individuals to incorporate many detailed circumstances and features into their predictions (i.e., internal model), leading to their expectations of facial expressions of emotion being more detailed and less generalizable or stereotypical than non-autistic individuals'.

In the present study we aimed to go beyond assessing facial expression predictions, additionally assessing the relevance of such predictions for social evaluation (likability ratings). In general, individuals tend to like people more when they display facial expressions that better match their predictions for a given emotionally evocative context (Chanes et al., 2018). Importantly, no differences were observed across groups in terms of the strength of this association, with autistic individuals using facial expression predictions to assess likability to the same extent as control individuals. These results show similarities with those of a recent study that found differences between autistic and control individuals when looking at accuracy levels in a facial emotion recognition task, but not when looking at a social functioning-related aspect. More specifically, while autistic individuals exhibited a less accurate retrospective emotion identification, a related social domain (specifically the authors evaluated empathy) appeared to be intact (Santiesteban et al., 2021). Thus, whereas autistic individuals seem to indeed differ from control individuals in their predictions about facial expressions, holding less stereotypical ones, they seem to use their predictions to a similar extent for social evaluation.

Our study is not without limitations. First, the autistic group included participants with co-occurring mental health conditions, and most were taking medication, something that should be taken into account in future studies. At the same time, we deemed it important to test a representative sample of individuals with autism, who do often show co-occurring mental health conditions. Second, all the included participants were adults without an intellectual developmental disorder, although recent studies show that up to a third of the autistic population present differences in intellectual development (Maenner et al., 2021). Thus, future studies may address these processes within a wider range of intellectual capabilities, aiming to better characterize the whole autistic spectrum. Also, in the present study we only asked whether their expectations were similar to the displayed expressions, but we did not assess what their expectations were specifically. Future work may characterize autistic individuals' expectations with regards to controls and assess whether differences exist with regards to clarity, level of detail, or other specific aspects. Moreover, we only assessed three emotion categories (happiness, sadness, and fear) and one social evaluation aspect (likability). Other emotions may be addressed in future studies in order to better explore and characterize the observed effects. Likewise, other ratings may be used in the future in order to explore social evaluation more thoroughly (e.g., trustworthiness, Chanes et al., 2018). Finally, we aimed to assess predictability as a whole, so how different aspects of the facial expression or emotion scenarios (e.g., facial configuration, intensity of the expression, specific cognitive aspects of image perception, narrative interpretations, affective response to narratives, etc.) may have contributed to the present results, remains to be further explored in the future.

Conclusion

The present study provides a new framework to understand how autistic individuals perceive social stimuli and evaluate others, revealing a deviation from stereotypicality in their predictions about others' emotion expressions. Importantly, though, the use of facial expression predictions in social evaluation (i.e., judgments of likability), appears to be intact among autistic individuals.

Funding

CRediT authorship contribution statement

Marta Robles: Formal analysis, Investigation, Data curation, Writing – original draft, Project administration. **Irene Ramos-Grille:** Investigation, Writing – review & editing. **Amaia Hervás:** Investigation, Writing – review & editing. **Enric Duran-Tauleria:** Investigation, Resources, Writing – review & editing. **Jordi Galiano-Landeira:** Writing – review & editing, Visualization. **Jolie B. Wormwood:** Writing – review & editing. **Christine M. Falter-Wagner:** Writing – review & editing, Funding acquisition. **Lorena Chanes:** Conceptualization, Methodology, Resources, Writing – original draft, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper

Acknowledgment

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders*. American Psychiatric Association. https://doi.org/10.1176/appi.books.9780890425596

American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders* (Fifth Edition). American Psychiatric Association. Text Revision (DSM-5-TR).

Barret, L. F. (2017). *How emotions are made: The secret life of the brain*. New York, NY: Houghton Mifflin Harcourt.

Barrett, L. F. (2017). The theory of constructed emotion: An active inference account of interoception and categorization. *Social Cognitive and Affective Neuroscience, 12*(1), 1–23.

Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest, 20*(1), 1–68. https://doi.org/10.1177/1529100619832930

Billieux, J., Rochat, L., Ceschi, G., Carré, A., Offerlin-Meyer, I., Defeldre, A. C., & Van der Linden, M. (2012). Validation of a short French version of the UPPS-P impulsive behavior scale. *Comprehensive Psychiatry, 53*(5), 609–615. https://doi.org/10.1016/j.comppsych.2011.09.001

Cannon, J., O'Brien, A. M., Bungert, L., & Sinha, P. (2021). Prediction in autism spectrum disorder: A systematic review of empirical evidence. *Autism Research, 14*(4), 604–630. https://doi.org/10.1002/AUR.2482

Chanes, L., Wormwood, J. B., Betz, N., & Barrett, L. F. (2018). Facial expression predictions as drivers of social perception. *Journal of Personality and Social Psychology, 114*(3), 380–396. https://doi.org/10.1037/pspa0000108

Draganov, M., Galiano-Landeira, J., Doruk Camsari, D., Ramírez, J.-E., Robles, M., & Chanes, L. (2023). Noninvasive modulation of predictive coding in humans: Causal

evidence for frequency-specific temporal dynamics. *Cerebral Cortex*. https://doi.org/10.1093/cercor/bhad127

Durán, J. I., & Fernández-Dols, J.-M. (2021). Do emotions result in their predicted facial expressions? A meta-analysis of studies on the co-occurrence of expression and emotion. *Emotion, 21*(7), 1550–1569. https://doi.org/10.1037/emo0001015 (Washington, D.C.).

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods, 12*(2), 121–138. https://doi.org/10.1037/1082-989X.12.2.121

Fernández-Dols, J.-M., & Crivelli, C. (2013). Emotion and expression: Naturalistic studies. *Emotion Review, 5*(1), 24–29. https://doi.org/10.1177/1754073912457229

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences, 360*(1456), 815–836. https://doi.org/10.1098/rstb.2005.1622

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

Friston, K. J., Lawson, R., & Frith, C. D. (2013). On hyperpriors and hypopriors: Comment on pellicano and burr. *Trends in Cognitive Sciences, 17*, 1. https://doi.org/10.1016/j.tics.2012.11.003

Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology, 85*(2), 348–362. https://doi.org/10.1037/0022-3514.85.2.348

Hoekstra, R. A., Vinkhuyzen, A. A. E., Wheelwright, S., Bartels, M., Boomsma, D. I., Baron-Cohen, S., & van der Sluis, S. (2011). The construction and validation of an abridged version of the autism-spectrum quotient (AQ-Short). *Journal of Autism and Developmental Disorders, 41*(5), 589–596. https://doi.org/10.1007/s10803-010-1073-0

Hoemann, K., Devlin, M., & Barrett, L. F. (2020). Comment: Emotions are abstract, conceptual categories that are learned by a predicting brain. *Emotion Review, 12*(4), 253–255.

JASP Team. (2022). JASP (Version 0.9.2) [Computer software].

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

Lozier, L. M., Vanmeter, J. W., & Marsh, A. A. (2014). Impairments in facial affect recognition associated with autism spectrum disorders: A meta-analysis. *Development and Psychopathology, 26*(4 Pt 1), 933–945. https://doi.org/10.1017/S0954579414000479

Lugo-Marín, J., Díez-Villoria, E., Magán-Maganto, M., Pérez-Méndez, L., Alviani, M., de la Fuente-Portero, J. A., & Canal-Bedia, R. (2019). Spanish validation of the autism quotient short form questionnaire for adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 49*(11), 4375–4389. https://doi.org/10.1007/s10803-019-04127-5

Maenner, M. J., Shaw, K. A., Bakian, A. V, Bilder, D. A., Durkin, M. S., Esler, A., & Cogswell, M. E. (2021). Prevalence and characteristics of autism spectrum disorder among children aged 8 years - autism and developmental disabilities monitoring network, 11 sites, United States, 2018. *Morbidity and Mortality Weekly Report, 70*(11), 1–16. https://doi.org/10.15585/mmwr.ss7011a1. Surveillance Summaries (Washington, D.C. : 2002).

Pellicano, E., & Burr, D. (2012). When the world becomes "too real": A Bayesian explanation of autistic perception. *Trends in Cognitive Sciences, 16*, 504–510. https://doi.org/10.1016/j.tics.2012.08.009

Quattrocki, E., & Friston, K. (2014). Autism, oxytocin and interoception. *Neuroscience and Biobehavioral Reviews, 47*, 410–430. https://doi.org/10.1016/j.neubiorev.2014.09.012

Ramos-Grille, I., Weyant, J., Wormwood, J. B., Robles, M., Vallès, V., Camprodon, J. A., & Chanes, L. (2022). Predictive processing in depression: Increased prediction error following negative valence contexts and influence of recent mood-congruent yet irrelevant experiences. *Journal of Affective Disorders, 311*, 8–16. https://doi.org/10.1016/j.jad.2022.05.030

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: SAGE Publications.

Santiesteban, I., Gibbard, C., Drucks, H., Clayton, N., Banissy, M. J., & Bird, G. (2021). Individuals with autism share others' emotions: Evidence from the continuous affective rating and empathic responses (CARER) task. *Journal of Autism and Developmental Disorders, 51*(2), 391–404. https://doi.org/10.1007/s10803-020-04535-y

Sinha, P., Kjelgaard, M. M., Gandhi, T. K., Tsourides, K., Cardinaux, A. L., Pantazis, D., & Held, R. M. (2014). Autism as a disorder of prediction. *Proceedings of the National Academy of Sciences of the United States of America, 111*(42), 15220–15225. https://doi.org/10.1073/pnas.1416797111

Smith, R., Badcock, P., & Friston, K. J. (2021). Recent advances in the application of predictive coding and active inference models within clinical neuroscience. *Psychiatry and Clinical Neurosciences, 75*(1), 3–13. https://doi.org/10.1111/pcn.13138

Tracy, J. L., & Randles, D. (2011). Four models of basic emotions: A review of Ekman and Cordaro, izard, Levenson, and Panksepp and watt. *Emotion Review, 3*, 397–405. https://doi.org/10.1177/1754073911410747

Uljarevic, M., & Hamilton, A. (2013). Recognition of emotions in autism: A formal meta-analysis. *Journal of Autism and Developmental Disorders, 43*(7), 1517–1526. https://doi.org/10.1007/s10803-012-1695-5

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*(4), 649–675. https://doi.org/10.1037/a0037665

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*(6), 1063–1070. https://doi.org/10.1037/0022-3514.54.6.1063

Wilson-Mendenhall, C. D., Barrett, L. F., & Barsalou, L. W. (2013). Neural evidence that human emotions share core affective properties. *Psychological Science, 24*(6), 947–956. https://doi.org/10.1177/0956797612464242

Yeung, M. K. (2022). A systematic review and meta-analysis of facial emotion recognition in autism spectrum disorder: The specificity of deficits and the role of task characteristics. *Neuroscience and Biobehavioral Reviews, 133*, Article 104518.

Yon, D., & Frith, C. D. (2021). Precision and the Bayesian brain. *Current Biology, 31*(17), R1026–R1032. https://doi.org/10.1016/j.cub.2021.07.044