Received: 4 April 2022 | Accepted: 23 June 2022

DOI: 10.1002/aur.2781

RESEARCH ARTICLE

Facing up to others' emotions: No evidence of autism-related deficits in metacognitive awareness of emotion recognition

Neil Brewer | **Carmen A. Lucas** | **Marie Antonia Georgopoulos** | **Robyn L. Young**

College of Education, Psychology & Social Work, Flinders University, Adelaide, Australia

Correspondence
Neil Brewer, College of Education, Psychology and Social Work, Flinders University, GPO Box 2100 Adelaide 5001, South Australia, Australia.
Email: neil.brewer@flinders.edu.au

Funding information
Australian Research Council, Grant/Award Number: DP190100162; Hamish Ramsay Fund

Abstract

Emotion recognition difficulties are considered to contribute to social-communicative problems for autistic individuals and awareness of such difficulties may be critical for the identification and pursuit of strategies that will mitigate their adverse effects. We examined metacognitive awareness of face emotion recognition responses in autistic ($N = 63$) and non-autistic ($N = 67$) adults across (a) static, dynamic and social face emotion stimuli, (b) free- and forced-report response formats, and (c) four different sets of the six "basic" and six "complex" emotions. Within-individual relationships between recognition accuracy and post-recognition confidence provided no indication that autistic individuals were poorer at discriminating correct from incorrect recognition responses than non-autistic individuals, although both groups exhibited marked inter-individual variability. Although the autistic group was less accurate and slower to recognize emotions, confidence-accuracy calibration analyses provided no evidence of reduced sensitivity on their part to fluctuations in their emotion recognition performance. Across variations in stimulus type, response format and emotion, increases in accuracy were associated with progressively higher confidence, with similar calibration curves for both groups. Calibration curves for both groups were, however, characterized by overconfidence at the higher confidence levels (i.e., overall accuracy less than the average confidence level), with the non-autistic group contributing more decisions with 90%–100% confidence. Comparisons of slow and fast responders provided no evidence of a "hard-easy" effect—the tendency to exhibit overconfidence during hard tasks and underconfidence during easy tasks—suggesting that autistic individuals' slower recognition responding may reflect a strategic difference rather than a processing speed limitation.

Lay Summary

It is generally considered that autistic individuals may have difficulty recognizing other people's facial emotions. However, little is known about their awareness of any emotion recognition difficulties they may experience. This study indicates that, although there is considerable individual variability, autistic adults were as sensitive to variations in the accuracy of their recognition of others' emotions as their non-autistic peers.

KEYWORDS
autistic adults, confidence-accuracy calibration, emotion recognition, metacognitive awareness

INTRODUCTION

The possibility that difficulties with social communication and interaction that are diagnostic features of autism (DSM-5; American Psychiatric Association, 2013) might be underpinned by problems recognizing facial expressions of others' emotions has attracted considerable research interest (e.g., Harms et al., 2010; Lozier et al., 2014; Nuske et al., 2013; Uljarević & Hamilton, 2013). It seems inevitable that difficulty recognizing others' emotions, or tardiness in doing so, will constrain understanding of how social interactions are unfolding and the capacity to respond appropriately. But other aspects of emotion processing may also have significant implications for the smoothness of social interactions. Awareness of one's limitations in recognizing others' emotions—metacognitive awareness—is also likely to be critical for the implementation of strategies that will mitigate any adverse impacts of emotion recognition difficulties. Yet surprisingly, researchers have paid scant attention to autistic individuals' awareness of any emotion recognition difficulties they may experience. The current research addresses this shortcoming.

In a recent study we examined speed and accuracy of recognition of face emotion expressions in autistic and non-autistic adults, incorporating (a) static photographs, dynamic moving images and social stimuli varying in the degree of contextual information, (b) free-report and multiple-choice response formats, and (c) 12 different emotions from the categories often referred to as "basic" or "complex" emotions (Georgopoulos et al., 2022). Emotion recognition accuracy—indicated by the level of agreement with the recognition responses of a normative sample (cf. Barrett et al., 2019)—and response latency data reported by Georgopoulos et al. (2022) indicate that task difficulty decreased progressively from static to dynamic to social stimulus presentations, and increased from multiple-choice to free-report responding and from basic to complex emotions. However, group differences remained robust across these conditions. Although there was considerable inter-individual variability in each group, for autistic individuals, recognition accuracy was lower (a relatively weak effect), response latencies were longer, and confidence in recognition responses was lower.

The emotion recognition data reported by Georgopoulos et al. (2022) derived from a larger project that also assessed metacognitive monitoring of emotion recognition responses and whether individuals could identify appropriate empathic responses to emotions displayed by others. This article focuses on the metacognitive monitoring component.

METACOGNITIVE AWARENESS OF EMOTION RECOGNITION

Studies of metacognitive awareness in autistic individuals in different task domains have produced mixed findings. Research with child and adolescent samples has reported lower metacognitive awareness in autistic than non-autistic individuals for general knowledge (Grainger et al., 2016; Williams et al., 2018) and mathematics tasks (Brosnan, Johnson, et al., 2016; but see Maras et al., 2019). Studies with adults have reported no group differences for a general knowledge task (Grainger et al., 2014a), episodic memory recall (Maras et al., 2020) or on post-decision wagering on the accuracy of perceptual discrimination judgments (Carpenter et al., 2019). However, poorer metacognitive awareness has been reported for autistic adults asked to provide "feeling of knowing" judgments (i.e., judgments of subsequent memory for previously unrecalled items) about cued recall and recognition of previously studied word pairs (Grainger et al., 2014b) or to interpret a range of mindreading, or Theory of Mind (ToM), scenarios (Brewer et al., 2022). Poor metacognition has also been implicated as a factor underlying autistic adults' longer latencies for post-decision wagering judgments about perceptual discrimination accuracy (Carpenter et al., 2019).

Studies of metacognitive awareness of emotion recognition in autistic individuals are rare and limited in scope. Several approaches can be used to evaluate metacognitive awareness. One is to examine whether "global" metacognition measures that probe individuals' perceptions of their awareness of or sensitivity to others' emotions using self-report questionnaires are related to objective performance on emotion recognition tasks. However, such measures generally fail to predict objective emotion recognition performance (e.g., Kelly & Metcalfe, 2011), potentially leading to the incorrect conclusion that people lack awareness of their discrimination of others' emotions.

A second approach is to elicit confidence judgments after each emotion recognition response and examine whether within-individual Goodman-Kruskal gamma (G) confidence-accuracy correlations indicate an ability to discriminate accurate from inaccurate responses (cf. Kelly & Metcalfe, 2011). This statistic, referred to as an index of resolution, identifies how well the confidence judgments discriminate correct from incorrect decisions and can range from $+1.0$ (perfect discrimination) to 0 (no association) and to -1.0 (perfect negative association).

Using this method, Sawyer et al. (2014) tested autistic and non-autistic adults using static images and a multiple-choice response format, with participants making confidence judgments after every response. Although emotion recognition performance was poorer for the autistic sample, individuals in both groups showed similar patterns of discrimination between accurate and inaccurate responses: G coefficients were 0.53 and 0.57 for autistic and non-autistic groups, respectively. In contrast, with participants aged 9–17 years, McMahon et al. (2016) reported poorer discrimination of accurate from inaccurate recognition responses of six basic emotions (static stimulus presentation) by the autistic than the non-autistic group ($G = 0.19$ vs. $G = 0.45$). It is impossible to determine whether the different findings reflect participants' developmental levels and whether the patterns will generalize across stimulus presentation types, response formats and emotions.

The emotion recognition database reported by Georgopoulos et al. (2022)—encompassing 340 observations per individual and more than 21,000 observations per group from samples of 63 autistic (21,420 trials) and 67 non-autistic (22,780 trials) adults—provides an excellent opportunity for examining metacognitive awareness within and between autistic and non-autistic groups. Although, given the potential contribution of co-occurring conditions, any differences between groups would not necessarily implicate autism-specific aspects of metacognitive awareness of emotion recognition, the data provide an opportunity to supplement the limited knowledge base on metacognitive awareness of emotion processing in autistic individuals.

Consistent with previous research on autistic individuals' metacognitive awareness, here we first examined within-individual confidence-accuracy *G* coefficients for autistic and non-autistic adults' emotion recognition responses for different types of stimulus presentation (static, dynamic, social), response format (free-report, multiple-choice), and sets of basic (afraid, angry, disgusted, happy, sad, surprised) and complex (ashamed, disappointed, frustrated, hurt, jealous, worried) emotions. These data inform understanding of group-level differences in the discrimination of accurate from inaccurate recognition responses (i.e., resolution) and of variability across individuals.

Although knowing that confidence levels are, on average, higher for accurate than inaccurate responses reflects one component of metacognitive awareness, it is not informative about whether confidence judgments are well calibrated with the probability that the response is accurate. For example, similar sized correlation coefficients may obtain when an individual provides (a) similar and very low confidence estimates for inaccurate responses and slightly higher but still relatively low estimates for accurate responses, (b) similar and relatively low confidence estimates for inaccurate responses and relatively high estimates for accurate responses, or (c) similar and moderately high confidence estimates for inaccurate responses and even higher estimates for accurate responses. More nuanced information about metacognitive awareness is provided by a third approach, confidence-accuracy calibration, that has been widely used across various decision-making domains (e.g., Baranski & Petrusic, 1994; Brewer & Wells, 2006; Cooke, 1906; Maras et al., 2020).

The calibration approach gives rise to a calibration curve plotting accuracy variations across the full range of confidence judgments (i.e., proportion correct for decisions made with 100% confidence, 90% confidence, etc.). It requires a sufficiently large number of observations per condition or group to provide stable estimates of accuracy at each confidence level, thereby indicating the sensitivity of adjustments in confidence to the range of variations in accuracy. Perfect calibration is indicated if 100% of decisions made with 100% confidence are accurate, 90% of decisions made with 90% confidence are accurate, and so on. However, if for example only 75% of decisions made with 100% confidence are accurate, 60% of decisions made with 80% confidence are accurate, and so on, it would indicate decision-making characterized by overconfidence. Conversely, underconfident decision-making would be suggested by the probability of an accurate response at each confidence level exceeding the confidence level (e.g., 80% of decisions made with 60% confidence are accurate). In contrast, the *G* coefficient indicates resolution: whether confidence is reliably higher for correct than incorrect responses.

Previously we summarized a limited body of evidence suggesting that, at least in some domains, autistic adults' metacognitive judgments may match those of non-autistic adults with respect to resolution. Robust evidence on the sensitivity of adults' metacognitive judgments, as provided by calibration analyses, is non-existent. To provide a comprehensive picture of metacognitive awareness of emotion recognition in autistic and non-autistic individuals, we focused on two issues. First, we used confidence judgments obtained after each recognition response to compute individual *G* correlations, focusing both on variability within groups and differences between groups to examine whether autistic and non-autistic individuals discriminated accurate from inaccurate recognition judgments.

Second, we derived confidence-accuracy calibration curves for the autistic and non-autistic groups across different types of stimulus presentation, response formats and emotion types. A robust finding across decision making domains is that calibration curves are characterized by increasing overconfidence as task difficulty increases—referred to as the hard-easy effect (e.g., Juslin et al., 2000). These patterns have been consistently reported in domains such as eyewitness identification (Palmer et al., 2013; Sauer et al., 2010), face recognition (Weber & Brewer, 2004), general knowledge (Baranski & Petrusic, 1995; Gigerenzer et al., 1991; Lichtenstein & Fischhoff, 1977), perceptual judgments (Baranski & Petrusic, 1994, 1995, 1998) and perspective taking (Brewer et al., 2022). The recognition accuracy and latency data reported by Georgopoulos et al. (2022) indicate that task difficulty decreased progressively from static to dynamic to social stimulus presentation, and increased from multiple-choice to free-report responding, and from basic to complex emotions. Calibration curves characterized by increasing overconfidence would, therefore, be expected to track the increasing difficulty levels indicated by the accuracy and latency data. Further, given Georgopoulos et al. (2022) reported lower accuracy for the autistic than the non-autistic group regardless of stimulus type, response format and emotion, the calibration patterns for the former group might be expected to be characterized by greater overconfidence.

Moreover, as Georgopoulos et al. (2022) argued, autistic individuals' longer recognition latencies might reflect either a fundamental difficulty processing emotions or perhaps a qualitatively different strategic approach—for example, exercising greater caution before finalizing decisions (cf. Pachella, 1974) or perhaps a disposition toward deliberative or effortful processing rather than the intuitive processing that some argue is more likely to be seen in non-autistic individuals (see Brosnan et al., 2017; Brosnan, Lewton, 2016—see, however, Taylor et al.'s (2022) compelling evidence against the latter position). Calibration curves characterized by greater overconfidence for the autistic than the non-autistic group would be expected if their longer latencies reflect a specific processing difficulty, but not if they reflect a different strategic approach. Similarly, if autistic individuals' longer latencies reflect greater processing difficulty, we might expect curves for slower autistic individuals to be characterized by greater overconfidence than those for faster individuals, consistent with the hard-easy effect.

We also conducted an exploratory examination of two issues: (1) Prior to commencing each trial-block for the different types of stimulus presentation, participants indicated how confident they were about recognizing others' emotions, thereby enabling a rudimentary group comparison of prospective awareness of emotion recognition ability (cf. Grainger et al., 2014b). (2) The availability of a ToM measure for the autistic participants from our database permitted a preliminary exploration of the relationship between metacognitive awareness and ToM. We present a brief discussion of the issue and the relevant data in Supplementary Materials (pp.13–15).

To summarize, we examined autistic and non-autistic individuals' metacognitive awareness of their emotion recognition performance under a variety of emotion stimulus and response conditions: free-report and multiple-choice recognition responses of emotions displayed in static, dynamic and social stimuli across multiple trials for 12 different emotions. Each response was followed by the participant indicating their confidence in the accuracy of that response. Given the limited prior research, the predictions we have foreshadowed thus far were, with one exception, tentative at best. For example, we made no directional predictions regarding how effectively autistic adults would discriminate correct from incorrect recognition responses (i.e., resolution), especially across the different stimulus presentation types, response formats and emotions. With respect to the calibration analyses, we rather confidently expected calibration curves for both groups to be characterized by (a) reduced overconfidence in parallel with the clear reductions in task difficulty from static to dynamic to social stimulus presentation, and (b) increased overconfidence paralleling the increases in difficulty from multiple-choice to free-report responding, and from basic to complex emotions. In contrast, we have outlined two possibilities for the calibration patterns for autistic and non-autistic groups. To the extent that autistic individuals' longer recognition latencies reflect a processing deficit, greater overconfidence would be expected, with this pattern even more marked for slower than faster responses. If, however, longer latencies simply reflect a different strategic approach such as more cautious responding, greater overconfidence for the autistic group would not be expected.

METHOD

We provide an abbreviated version of the Method section of Georgopoulos et al. (2022).

Participants

The autistic sample comprised 63 participants (17 female), aged 18–66 years ($M = 31.1$ $SD = 13$), with a diagnosis of autism spectrum disorder based on the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR [APA, 2000] or DSM-5 [APA, 2013]). Wechsler Abbreviated Scale of Intelligence-Second Edition (WASI-II; Wechsler, 2011) Verbal Comprehension Index (VCI) scores spanned 85–143 ($M = 104.44$, $SD = 12.71$, 95% CI [101.30, 107.58]). The non-autistic sample included 67 individuals (47 female), aged 18–65 years ($M = 23.8$ $SD = 8.9$) with VCI scores ranging from 85 to 136 ($M = 106.57$, $SD = 11.44$, 95% CI [103.83, 109.31]). The groups did not differ significantly in VCI, t (128) = 1.00, $p = 0.318$, $d = 0.18$, d 95% CI (−0.17, 0.52), but the non-autistic group was significantly younger, t (109.26) = 3.73, $p < 0.001$, $d = 0.66$, d 95% CI (0.31, 1.01).

Design

The design was a 2 (Group: autistic, non-autistic) × 2 (Response Format: free-report, multiple-choice) × 3 (Stimulus Type: static, dynamic, social) × 12 (Emotion: afraid, angry, ashamed, disappointed, disgusted, frustrated, happy, hurt, jealous, sad, surprised, worried) mixed design, with stimulus type, response format and emotion as within-subjects' factors.

Materials

Stimuli were obtained from the EU-Emotion Stimulus Set database (O'Reilly et al., 2012, 2016)[1]; we used the high intensity versions of 12 emotions, displayed as static images, dynamic video clips (a video of the person making the facial expression), and contextual social scenes, without vocalizations. Stimuli were delivered on a 15-inch Apple Macbook Pro; participants used a USB/wireless mouse to indicate their responses.

Response format

Participants viewed each image or short clip and typed a single-word free-report response to indicate the emotion the target stimulus was feeling. After providing a confidence rating for the free-report response, they selected the emotion the target was feeling from four multiple-choice options, followed again by a confidence rating. Multiple-choice options for each item included the target emotion and three foils randomly selected from a pool of 60 options that included the other 11 target emotions and 49 other emotions that were not too similar to the target emotions

---

[1] Bona fide researchers may access the stimuli via the database manager, H. O'Reilly of the ASC- Inclusion Project: Autism Research Centre, University of Cambridge, UK. Email: heo24@medschl.cam.ac.uk

Stimulus type

Trials for each stimulus type were completed in randomized order within a block, with presentation order of the blocks counterbalanced. The static task presented static photos of a facial configuration (from the shoulders up) whereas the dynamic task presented short video clips of a person (shoulders up) moving their face into configurations depicting the target emotions. For the static and dynamic stimulus types there were four trials for each emotion (i.e., 48 trials for each stimulus type). In the social stimulus task, there were 4–7 (*Mdn* = 7)[2] trials for each emotion; participants were presented with video clips of an interaction between two people that provided important contextual visual information, with participants providing a free-report and then a multiple-choice response indicating the emotion of the person in the interaction specified as the target individual.

The following URLs provide two examples of each task:

- Static: https://qualtrics.flinders.edu.au/jfe/form/SV_9pqp0YDttUmdts1
- Dynamic: https://qualtrics.flinders.edu.au/jfe/form/SV_8vafpqESU40Iyu9
- Social: https://qualtrics.flinders.edu.au/jfe/form/SV_9oad9qroRmS3yrr

Emotion

The 12 emotions used were the six 'basic' emotions (afraid, angry, disgusted, happy, sad, surprised) and six emotions (ashamed, disappointed, frustrated, hurt, jealous, worried) typically classified as "complex" emotions.

Measures

Recognition accuracy

Assessments of recognition accuracy relied on independent observers' subjective judgments to index normative interpretations of the emotional expressions rather than on some objective index. The procedures used for classifying the "accurate" multiple-choice response for each face emotion stimulus are described in O'Reilly et al. (2016).

Free-report accuracy: Coding of agreement with normative responses to emotions

Free-report responses that matched the appropriate emotion term or a synonym obtained from online thesaurus or dictionary platforms were coded as matching the normative response (i.e., accurate). Three judges scored non-synonym responses as 3 (strict: meaning the same as normative response), 2 (lax: similar meaning, but not exactly the same), 1 (boundary: plausible alternative but not really a synonym) or 0 (incorrect: not at all like the emotion). All coding classifications are available at https://osf.io/ndbfs/. For example, scores of 1, 3, 0, and 2 on the 4 trials for the static presentations of "angry" summed to 6 of a possible 12, producing an accuracy score of 50%. Inter-rater reliability indexed by Cohen's kappa was 0.89 and 0.83 for two independent raters' coding of two large subsets of free report responses (see Georgopoulos et al., 2022).

To calculate *G* coefficients for strict coding of free-report responses, responses scored as 3 were recoded as 1 (correct) and all other scores were recoded as 0 (incorrect). For lax coding, responses scored as 3 or 2 were recoded as 1 (correct) and scores of 1 and 0 were recoded as 0 (incorrect).

For multiple-choice responses, selection of the target emotion (the normative or accurate recognition response) on each trial = 1, and a foil selection = 0. For example, selecting the target response on 3 of the 4 trials for the static presentation of "angry" resulted in an accuracy score of 75%.

Latency and confidence

Recognition response latency was recorded to the nearest 0.01 s and confidence was indicated on an 11-point scale ranging from 0% to 100% confident that the response was accurate.

Procedure

At the beginning of the static, dynamic and social tasks, the experimenter read aloud the instructions on the laptop screen and ensured participants understood task requirements. After providing demographic information, participants rated their confidence in their ability to accurately recognize others' emotions, completed two practice trials for that task and were instructed to respond as quickly and accurately as possible, with the latter instruction appearing on the screen prior to each trial.

RESULTS

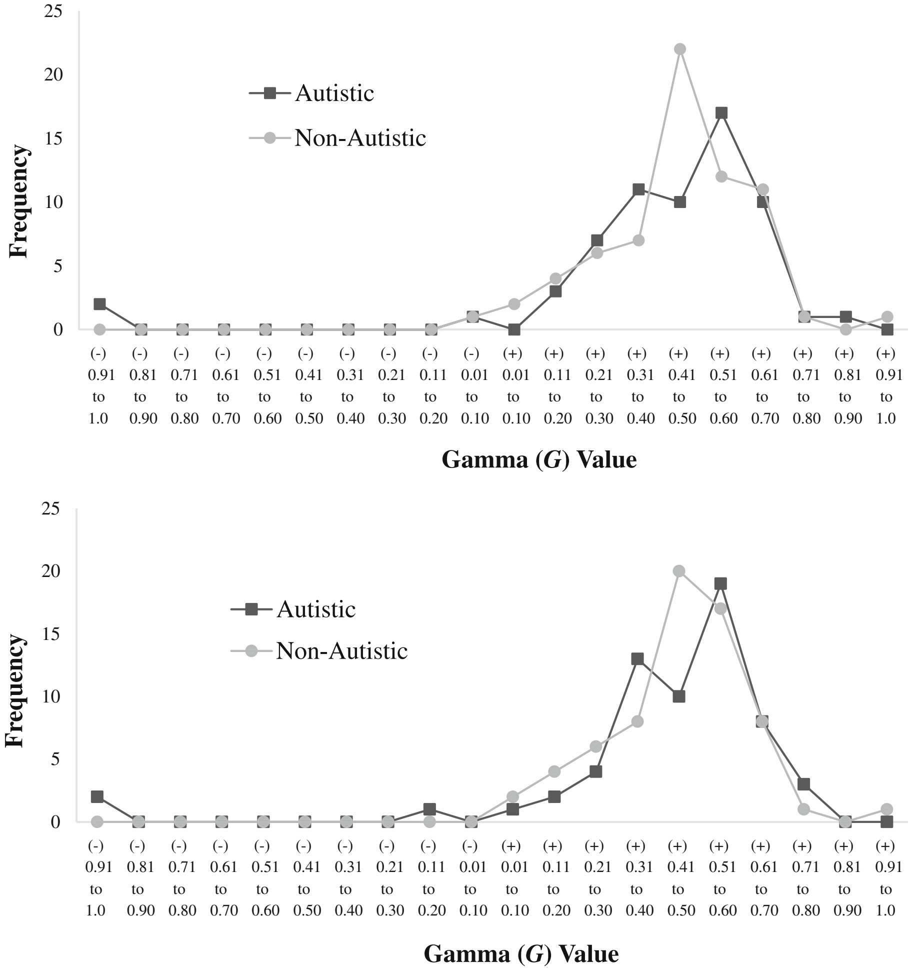
Within-individual confidence-accuracy relationships

Confidence-accuracy *G* coefficients were calculated for each participant, across all trials and for response format, stimulus type and emotion subsets using both strict and

---

[2] More trials were used in the social stimulus task where possible to provide a larger database for the phase of the broader project concerned with empathic reactions to others' emotions.

FIGURE 1 Distributions of the confidence-accuracy gamma (G) coefficients calculated from all responses (i.e., multiple choice and free-report) with strict (upper panel) and lax (lower panel) free-report coding



lax accuracy coding for free-report responses. Table S1 shows the range, mean and median *G* coefficients for all conditions and inferential group comparisons. For these analyses, the 12 emotions were collapsed into two categories: basic and complex. For both groups, coefficients were generally of similar magnitude across conditions.[3] For example, for free-report trials (strict coding), *G* ranged from −0.04 to 0.84 ($M = 0.43$, $SD = 0.16$, $Mdn = 0.43$) within the autistic group and from −0.05 to 0.72 ($M = 0.40$, $SD = 0.16$, $Mdn = 0.40$) in the non-autistic group, $t\,(123) = 1.01$, $p = 0.315$, $d = 0.18$, d 95% CI (−0.17, 0.52).

Figure 1 highlights the similar *G* distributions calculated from all trials for both groups under free-report strict (upper panel) and lax (lower panel) coding. These indicate that both groups were, at least to some degree, aware of when they had recognized the displayed emotions. There was, however, marked inter-individual variability within each group, with some individuals good at discriminating correct from incorrect responses whereas others were not.

We also investigated whether individuals' awareness of their ability to discriminate others' emotions was associated with their recognition of emotions. Table S2 shows the correlations between accuracy and confidence-accuracy *G* correlation coefficients (overall and by condition). For free-report (strict coding), for example, $r_s = 0.32$, $p = 0.014$, and $r_s = -0.05$, $p = 0.717$, for autistic and non-autistic individuals, respectively, $z = 2.08$, $p = 0.019$. This pattern—moderate and significant correlations for autistic individuals versus often weaker and non-significant correlations for non-autistic individuals—was observed across many of the conditions (see Table S2).

Confidence-accuracy calibration

Confidence data were collapsed into five categories (i.e., 0%–20%, 30%–40%, 50%–60%, 70%–80%, 90%–100%) to maximize stability of estimates in each confidence category. The proportion of accurate recognition responses in each category was then plotted against the weighted mean confidence for that category. For the free-report condition, we applied the strict criterion for accuracy scores. The resultant calibration curves for autistic and non-autistic groups in the different conditions are

---

[3] As shown in Table S1, the non-parametric contrasts for the dynamic stimulus condition revealed significantly higher coefficients for autistic than non-autistic individuals but the effects were weak.

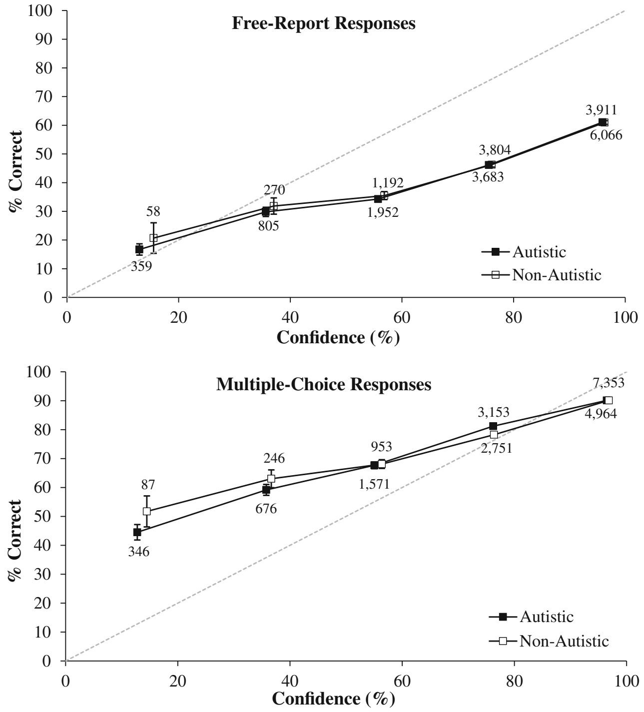


FIGURE 2 Confidence-accuracy calibration curves for autistic and non-autistic groups for the two response formats (free-report, multiple-choice). Number of data points in each confidence category is shown. Dotted line represents perfect calibration. Error bars represent standard error. For free-report responses, $C = 0.08$, $O/U = 0.27$ (autistic) and $C = 0.10$, $O/U = 0.31$ (non-autistic); for multiple-choice responses, $C = 0.01$, $O/U = -0.03$ (autistic) and $C = 0.01$, $O/U = 0.02$ (non-autistic).

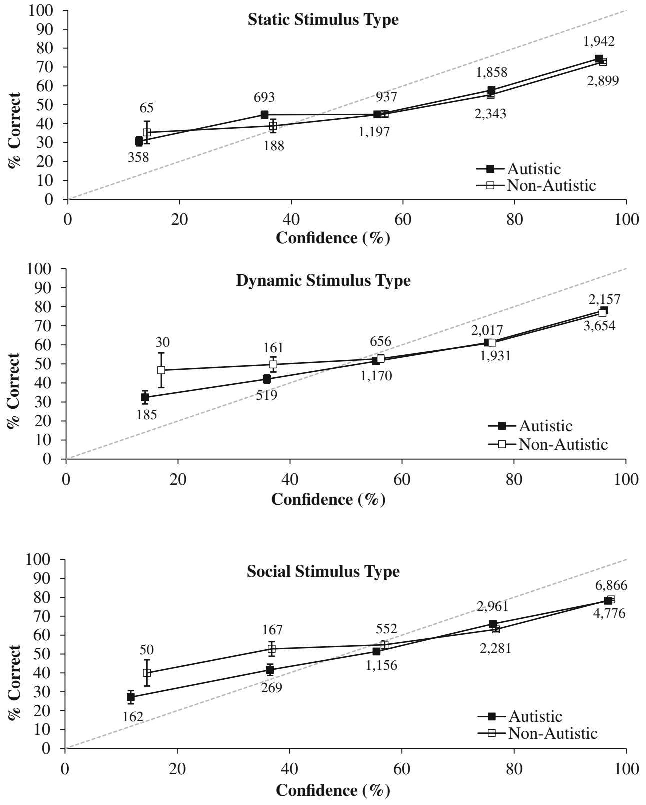


FIGURE 3 Confidence-accuracy calibration curves for autistic and non-autistic groups for the three stimulus types (i.e., static, dynamic and social). For static stimuli, $C = 0.03$, $O/U = 0.12$ (autistic) and $C = 0.04$, $O/U = 0.19$ (non-autistic); for dynamic stimuli, $C = 0.02$, $O/U = 0.11$ (autistic) and $C = 0.03$, $O/U = 0.15$ (non-autistic); for social stimuli, $C = 0.02$, $O/U = 0.13$ (autistic) and $C = 0.03$, $O/U = 0.16$ (non-autistic).

displayed in Figure 2 (response format), Figure 3 (stimulus type), and Figure 4 (emotion).

We also examined two statistics that assist interpretation of the calibration curves: the *C* (calibration) statistic and the *O*/*U* (over/underconfidence) statistic (see Baranski & Petrusic, 1994; Lichtenstein et al., 1982). The *C* statistic can vary from 0 to 1, with zero indicating perfect calibration. The *O*/*U* statistic varies from +1 to −1, with more marked over- and underconfidence indicated by larger positive and negative scores, respectively. The *O*/*U* statistic denotes whether participants' average confidence is greater than (indicating overconfidence) or less than (underconfidence) overall accuracy. To calculate *C*, (1) each confidence rating is assigned to a class interval; (2) the difference between the mean confidence level and the proportion of correct responses for each class interval is obtained; (3) the squared differences are multiplied by the number of observations in the interval; and (4) the step (3) outcomes are summed across class intervals and divided by the total number of observations. The *O*/*U* statistic is calculated in the same way as the *C* statistic, except that the differences are not squared, and thus it is equivalent to mean confidence minus overall proportion correct. Two calibration curves could run parallel to the ideal calibration curve and yield similar *C* statistics, but with one lying below the ideal (yielding a positive *O*/*U* statistic) and the other above it (negative *O*/*U* statistic). The former would indicate overconfidence and the latter underconfidence. The *C* and *O*/*U* statistics were derived for each individual by each condition, and overall, and then analyzed with parametric tests (cf. Weber & Brewer, 2004, for an application with a face recognition paradigm).

We highlight six features of the calibration curves. First, in all conditions, recognition accuracy increased with increases in confidence, indicating that participants were monitoring fluctuations in the accuracy of their responses with some degree of sensitivity. Second, none of the calibration curves tracked the dotted line indicating perfect calibration. Third, in the upper sections of the scale, in all conditions, the curves were characterized by overconfidence (i.e., accuracy was lower than the corresponding confidence level), indicating participants' failure to lower confidence estimates sufficiently to align with accuracy. As expected, the degree of overconfidence was more pronounced in those conditions where accuracy reflected greater task difficulty. For example, accuracy at confidence levels of 90–100% was (a) lower for free-report than multiple-

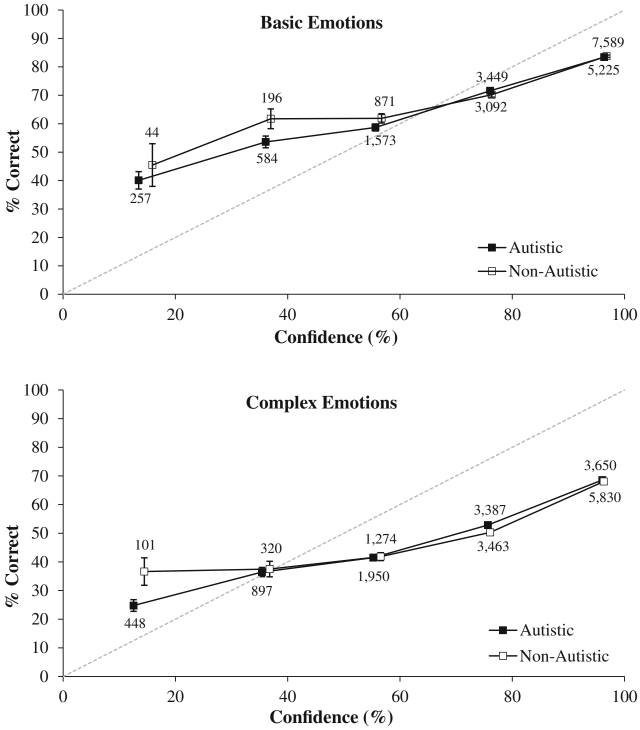


FIGURE 4 Confidence-accuracy calibration curves for autistic and non-autistic groups for basic and complex emotions. For basic emotions, $C = 0.01$, $O/U = 0.06$ (autistic) and $C = 0.01$, $O/U = 0.09$ (non-autistic); for complex emotions, $C = 0.05$, $O/U = 0.19$ (autistic) and $C = 0.07$, $O/U = 0.25$ (non-autistic).

choice responses (Figure 2), (b) lowest for static, followed by dynamic and then social stimuli (Figure 3), and (c) lower for complex than basic emotions (Figure 4). Fourth, in the lower sections of the scale, the curves were characterized by underconfidence (except for the clearly more difficult free-report condition), indicating participants lowered their confidence estimates more than was necessary to align with the associated accuracy levels. Fifth, apart from the non-autistic group being characterized (under some conditions) by more underconfidence when using the lower section of the scale—although the associated number of observations was relatively small—the calibration curves for the two groups were virtually overlapping. Sixth, note, however, that at the 90%–100% confidence levels, non-autistic individuals contributed many more observations than autistic individuals.

This descriptive overview of the calibration curves is supported by formal analyses of the C and O/U statistics for each participant. These statistics, by group and condition, are presented in Table 1. Three separate 2 (Group: autistic, non-autistic) × 3 (Stimulus Type: static, dynamic, social scenes), 2 (Group: autistic, non-autistic) × 2 (Response Format: multiple-choice, free-report) and 2 (Group: autistic, non-autistic) × 2 (Emotion: basic, complex) ANOVAs were run on the C and O/U statistics, with stimulus type, response format and emotion as within-subjects factors.[4] As the data violated one or more assumptions of parametric ANOVA, non-parametric versions of each analysis were also conducted using the WRS2 package in R (version 1.1–3; Mair & Wilcox, 2019). Given the parallels between the results of the parametric (Table 2) and non-parametric analyses (Table S3), we focus on the parametric results but note some minor discrepancies below, with reference to (Table S3).

There were effects of response format, stimulus type and emotion on both C ($Fs \geq 19.34$, $ps < 0.001$) and O/U ($Fs \geq 3.34$, $ps \leq 0.040$), with strong effects of response format and emotion on both C ($\eta_G^2 = 0.411$ and 0.239, respectively) and O/U ($\eta_G^2 = 0.622$ and 0.287, respectively) but only weak effects of stimulus type on C ($\eta_G^2 = 0.048$) and O/U ($\eta_G^2 = 0.007$). Consistent with the calibration curves, the descriptive statistics indicate poorer calibration and greater overconfidence for free-report than multiple-choice responses, $C = 0.11$ ($SD = 0.07$) and $O/U = 0.29$ ($SD = 0.12$) versus $C = 0.03$ ($SD = 0.02$) and $O/U = -0.00$ ($SD = 0.11$), and for complex than basic emotions, $C = 0.08$ ($SD = 0.06$) and $O/U = 0.22$ ($SD = 0.12$) versus, $C = 0.03$ ($SD = 0.02$), $O/U = 0.07$ ($SD = 0.11$). Calibration was better and overconfidence was less in the dynamic, $C = 0.05$ ($SD = 0.04$) and $O/U = 0.13$ ($SD = 0.13$), and social conditions, $C = 0.04$ ($SD = 0.04$) and $O/U = 0.14$ ($SD = 0.11$) when compared with the static trials, $C = 0.07$ ($SD = 0.06$) and $O/U = 0.16$ ($SD = 0.15$).

None of the analyses showed an effect of group on C ($Fs \leq 0.53$, $ps \geq 0.468$) and the effect sizes were trivial ($\eta_G^2 \leq 0.003$). There were no significant parametric analysis interaction effects involving group on C ($Fs \leq 3.93$, $ps \geq 0.050$, $\eta_G^2 \leq 0.009$), and in each case the effect sizes were tiny. The nonparametric analysis produced significant group × response format and group × emotion interactions, both reflecting larger, but still relatively small, differences in C for the non-autistic than the non-autistic group (see Table S3).

Conversely, as shown in Table 2, a relatively weak effect (η_G^2 ranged from 0.03 to 0.04) of group on O/U was observed in all analyses ($Fs \geq 5.53$, $ps \leq 0.020$), with non-autistic participants ($O/U = 0.17$, $SDs = 0.12$–0.13) exhibiting greater overconfidence than autistic participants ($O/U = 0.12$, $SDs = 0.09$–0.10) overall. There were no significant interaction effects on O/U ($Fs \leq 2.46$, $ps \geq 0.091$), with effect sizes ranging from weak ($\eta_G^2 = 0.037$) to very weak ($\eta_G^2 \leq 0.005$).

Although the effects of group on C (i.e., across subsets) and many interaction effects involving group on both C and O/U were non-significant and characterized by very small effect sizes, we conducted Bayesian analyses which confirmed the extent to which these reflected evidence in favor of the null. For each analysis, the best-fitting model did not include the predictor(s) associated with non-significant frequentist results. Evidence for the

---

[4] Separate ANOVAs, with data collapsed across the other conditions, were conducted to ensure stable estimates of C and O/U for each individual.

TABLE 1 Mean (and SD) *C* and *O*/*U* statistics for autistic and non-autistic individuals by condition

	Group			
Condition	**Autistic**		**Non-autistic**	
Response format	*C*	*O*/*U*	*C*	*O*/*U*
Free-report	0.11 (0.08)	0.27 (0.14)	0.12 (0.06)	0.31 (0.10)
Multiple-choice	0.03 (0.03)	−0.03 (0.12)	0.02 (0.02)	0.02 (0.10)
Stimulus type				
Static	0.07 (0.06)	0.12 (0.17)	0.07 (0.06)	0.19 (0.13)
Dynamic	0.05 (0.04)	0.11 (0.14)	0.05 (0.04)	0.15 (0.11)
Social	0.05 (0.04)	0.13 (0.13)	0.04 (0.02)	0.16 (0.09)
Emotion				
Basic	0.03 (0.03)	0.06 (0.12)	0.03 (0.02)	0.09 (0.10)
Complex	0.07 (0.07)	0.19 (0.14)	0.09 (0.06)	0.25 (0.11)

TABLE 2 Results of parametric ANOVAs examining the effects of group and response format, stimulus type, and emotion on *C* and *O*/*U* statistics

	C						*O*/*U*					
Predictor	*df*	**Mean square**	*F*	*p*	η_p^2	η_G^2	*df*	**Mean square**	*F*	*p*	η_p^2	η_G^2
Group	1	<0.001	0.01	0.944	<0.001	<0.001	1	0.135	5.70	0.018	0.043	0.038
Error	128	0.003					128	0.024				
Response format	1	0.486	184.72	<0.001	0.591	0.411	1	5.62	1932.33	<0.001	0.938	0.622
Group × response format	1	0.006	2.30	0.132	0.018	0.009	1	<0.001	0.29	0.589	0.002	<0.001
Error	128	0.003					128	0.003				
Group	1	0.001	0.28	0.598	0.002	0.001	1	0.225	6.03	0.015	0.045	0.034
Error	128	0.004					128	0.037				
Stimulus type[a]	1.68	0.024	19.34	<0.001	0.131	0.048	1.88	0.024	3.34	0.040	0.025	0.007
Group × stimulus type	1.68	0.001	0.59	0.524	0.005	0.002	0.188	0.017	2.46	0.091	0.019	0.005
Error	214.41	0.001					240.84	0.007				
Group	1	0.002	0.53	0.468	0.004	0.003	1	0.131	5.53	0.020	0.041	0.037
Error	128	0.003					128	0.024				
Emotion	1	0.170	131.07	<0.001	0.506	0.239	1	1.37	482.83	<0.001	0.790	0.287
Group × emotion	1	0.005	3.93	0.050	0.030	0.009	1	0.005	1.83	0.179	0.041	0.037
Error	128	0.001					128	0.003				

[a] Post-hoc tests indicated significant differences in *C* between the static condition and both the dynamic and social conditions and a significant difference in *O*/*U* between the static and dynamic conditions.

inclusion of these predictors ranged from anecdotal evidence for inclusion to strong evidence for non-inclusion (i.e., BF_{incl} values between 1.44 and 0.07). The full Bayesian results are summarized in Tables S4–S6.

In sum, despite lower recognition accuracy, longer latencies and lower confidence suggesting greater processing difficulty for autistic than non-autistic individuals (see Georgopoulos et al., 2022), the calibration curves for the autistic group did not indicate the poorer calibration or more marked overconfidence typically associated with greater task difficulty. Indeed, the non-autistic group contributed many more (though obviously a similar proportion of) highly confident incorrect responses. Moreover, when we created calibration curves comparing the fastest and slowest free-report trials (a median split) for autistic and non-autistic participants (see Figure S1), there was no suggestion that the calibration curve for the slower autistic individuals—those who might be presumed to have experienced greater processing difficulty—was characterized by greater overconfidence than that of the faster autistic individuals. Rather, overconfidence was less pronounced, and calibration closer to zero, for the slower ($C = 0.067$ and $O/U = 0.237$) than the faster ($C = 0.102$ and $O/U = 0.304$) autistic individuals.

Given the age difference between groups, we examined metacognitive awareness of younger and older autistic participants to check whether the absence of group differences might be due to the greater proportion of older individuals in the autistic group. Three lines of analyses ruled out this possibility. First, the correlations between age and the G, C and O/U statistics were all negligible. The respective coefficients were (i) $r(59) = -0.095$, $p = 0.465$, $r_s(59) = -0.026$, $p = 0.483$, $G = -0.017$, $p = 0.843$, for G (G calculated across all trials with strict FR coding), (ii) $r(61) = 0.033$, $p = 0.795$, $r_s(61) = 0.085$, $p = 0.505$, $G = 0.064$, $p = 0.475$, for C, and (iii) $r(61) = -0.063$, $p = 0.624$, $r_s(61) = -0.063$, $p = 0.625$, $G = -0.041$, $p = 0.649$, for O/U. Second, neither G coefficients, C statistics nor O/U statistics differed significantly for participants in the younger and older sub-groups based on a median split at 26 years (see Table S7). Third, the C and O/U statistics were consistent with the younger and older groups' almost overlapping calibration curves (see Figure S2). Together, these patterns fail to highlight any noteworthy effect of age on the patterns of metacognitive awareness for the two groups.

Exploratory analyses

Pre-task confidence

Autistic individuals were significantly less confident pre-task than non-autistic individuals (overall $M = 68.47$, $SD = 15.02$ vs. $M = 75.67$, $SD = 12.30$), $t(119.98) = -2.98$, $p = 0.003$, $d = 0.53$, 95% CI (0.18, 0.88). Lower pre-task confidence was associated with longer response latencies in autistic, $r_s(61) = -0.30$, $p = 0.019$, but not non-autistic individuals, $r_s(65) = -0.15$, $p = 0.234$ (although the difference between these coefficients was not significant, $z = 0.88$, $p = 0.189$). However, there was no suggestion that lower pre-task confidence was associated with lower overconfidence in either autistic $r_s(61) = 0.13$, $p = 0.307$, or non-autistic individuals, $r_s(65) = 0.01$, $p = 0.926$.

DISCUSSION

Our examination of metacognitive awareness of emotion recognition performance in autistic and non-autistic adults across various stimulus and response conditions incorporated (a) confidence-accuracy G coefficients, reflecting the ability to discriminate accurate from inaccurate responses, (b) C and O/U calibration statistics for individual participants and group level calibration data, indicating the sensitivity of post-decision adjustments in confidence to fluctuations in recognition accuracy—with all measures based on large numbers of observations.

Within-individual relationships between confidence and accuracy provided no indication that, at the group level, the autistic individuals were any less able to discriminate correct from incorrect recognition responses than non-autistic individuals, although inter-individual variability was marked within each group. Further, the confidence-accuracy calibration curves and statistics (supported by the Bayesian analyses) demonstrate quite similar degrees of sensitivity to performance fluctuations from both groups. Across variations in response format, stimulus type and emotion, higher accuracy was associated with greater confidence, with virtually identical calibration curves for both groups. Although both groups were characterized by overconfidence at the highest confidence levels, this pattern was more pronounced for non-autistic individuals.

As indicated previously, calibration curves are typically characterized by greater overconfidence as task difficulty increases. The calibration data—where the autistic group showed less overconfidence than the non-autistic group—suggest, therefore, that autistic individuals' longer response latencies may not have reflected greater task difficulty, but rather a difference of a strategic nature such as exercising greater caution before responding. Two data patterns are consistent with this possibility. First, the negative correlation between pre-task confidence and latency in the autistic group suggests the possibility that lower pre-task confidence translated into more cautious responding. Second, when we compared calibration curves for the slowest and fastest autistic individuals, the former curve was characterized by less overconfidence, a pattern that is inconsistent with an interpretation that their slowness reflecting greater task difficulty.

Nevertheless, further research is needed to provide direct confirmation of that conclusion. One approach might be to examine the effect on latency of providing post-trial (and possibly false positive) feedback designed to boost self-efficacy. If autistic individuals' longer latencies simply reflect caution, boosting belief in their social–emotional competence might lead to them adopting a less conservative approach. Another approach, outlined by Georgopoulos et al. (2022), could involve using limited exposure durations or response time deadlines to constrain more cautious processing (cf. Brewer & Smith, 1990; Tracy et al., 2011). If autistic-nonautistic differences in emotion recognition accuracy increased with progressively shorter exposure durations or response deadlines, it would suggest that decoding of face emotions is more difficult for autistic individuals. But if similar and asymptotic accuracy levels were maintained by both groups, the exercise of greater caution (or maybe the adoption of some form of more effortful strategic approach) on the part of autistic individuals would be suggested.

Should future research confirm that the lower overconfidence of autistic individuals does indeed reflect a particularly cautious approach, perhaps indicating an autism-related strength, what might be some of the implications? The answer would likely depend on whether greater caution is confined to the emotion recognition paradigm or also characterizes ‘real world’ emotion processing. If the latter, there may be negative consequences: for example, delayed decoding of others’ emotions, followed by delayed reactions, might be interpreted as a sign of disinterest or lack of empathy. Moreover, although a cautious approach may ultimately prove to be unnecessary for autistic individuals to achieve optimum accuracy, it may be one with which they are most comfortable. Accordingly, examination of individual’s affective reactions when processing or response times are constrained would be a worthwhile component of such investigations.

When reviewing previous studies in the Introduction, we noted that evidence for autistic-nonautistic group differences in metacognitive awareness appears largely (e.g., Grainger et al., 2016; Williams et al., 2018)—although not exclusively (e.g., Brewer et al., 2022; Grainger et al., 2014b)—confined to child and adolescent samples. Perhaps sustained exposure to face emotion stimuli over many years provides whatever critical experiences are necessary for the development of metacognitive awareness in this domain, even if the individual’s autistic characteristics may have led to diminished engagement with social stimuli. Answers to such questions will minimally require cross-sectional studies of child, adolescent, and adult samples and, ideally, longitudinal analysis.

An unexpected finding was that metacognitive awareness appeared to be more strongly related to emotion recognition performance in autistic than non-autistic individuals. Perhaps there are individuals who are aware that they are performing well because they have learned through exposure to intervention to identify, and focus their attention on, those critical perceptual cues used to facilitate recognition. Alternatively, perhaps “good recognizers” have become sensitive to cues such as perceived ease of processing or feelings of automaticity that help them infer whether their individual responses are likely to be accurate (cf. Koriat & Ackerman, 2010; Semmler et al., 2004). However, such explanations beg the question as to why a similar metacognition–performance relationship was not apparent in non-autistic individuals, especially given the similar distributions of within-individual *G* correlations. As these data patterns are only correlational, we can only speculate. It seems unlikely that only autistic individuals would be sensitive to cues that might lead to inferences of accurate recognition. But it has been proposed that autistic and non-autistic individuals differ in terms of perceptual processing, with the former more likely to be characterized by a local (cf. a global) processing style that is advantageous for certain perceptual discriminations—a view captured in weak central coherence and enhanced perceptual functioning theories of autism (e.g., Frith & Happé, 1994; Mottron et al., 2006). Thus, perhaps autistic individuals who effectively combine a local processing style with accurate inferences about when they are performing well are better suited than non-autistic individuals to analyze and exploit whatever fine perceptual details they have been using for recognition.

We highlight four broader questions that merit examination in future research. First, for those autistic and non-autistic individuals who were relatively poor at discriminating accurate from inaccurate recognition responses, the implications for the effectiveness of their interpersonal interactions warrant investigation. Although it may be tempting to infer that a lack of awareness of one’s capacity to decode others’ expressions will undermine the effectiveness of social exchanges, explicit investigation of such relationships is required.

Second, would the patterns observed for autistic and non-autistic individuals be replicated in the more complex and often multi-person, interpersonal interactions encountered in daily life? And, in such interactions, how might the involvement of other cues such as those provided by an interaction partner’s tone of voice and body gestures shape the individual’s evaluations of their interpretations of others’ emotions?

Third, our data provide no indication regarding whether either autistic or non-autistic individuals are likely to be responsive to any difficulties in emotion recognition of which they become aware and nor do they signal how they might respond. Knowing one’s limitations does not necessarily translate into adaptive steps to address those limitations. In other words, metacognitive limitations may exist at a superordinate level with, for example, some individuals not being aware of the potential importance of adjusting to any limitations in order to increase the effectiveness of future interactions. Radically different research paradigms will be required to address such questions.

Fourth, there has been considerable recent interest in the respective contributions of autism and alexithymia (i.e., a difficulty identifying and describing one’s own emotions) to the processing of emotions, with alexithymia sometimes identified as the more important predictor of emotion recognition difficulties (e.g., Cook et al., 2013). Note, however, Keating et al.’s (2022) finding that when non-static emotion stimuli were used, alexithymia did not predict emotion recognition accuracy. The contribution of alexithymia to the metacognitive awareness of participants in our study is obviously unknown and further research into this issue is warranted. However, we offer here some brief speculation about potentially important considerations in such research.

Although a recent meta-analysis (Huggins, Donnan, et al., 2021) clearly demonstrated poorer emotional self-awareness in autistic individuals, and especially in adults, Huggins et al. note the heavy reliance on self-report measures such as the Toronto Alexithymia Scale (Bagby et al., 1994). They suggest that such self-reports may simply reflect autistic individuals’ underestimation of their own competencies in this domain. Indeed, when emotional self-awareness was measured with a behavioral rather than a self-report measure, individuals with high autistics traits underestimated their emotional self-awareness whereas those with low autistic traits did the opposite (Huggins et al., 2021). Huggins et al., (2021) suggest that autistic individuals' poorer emotional self-awareness as measured with self-report instruments may simply reflect a lack of confidence in their own abilities.

Against this background, several of our findings are noteworthy. First, the autistic group showed lower prospective confidence than the non-autistic group on the emotion recognition task. Second, prospective confidence and response latency were negatively correlated in the autistic group, consistent with the possibility that lower prospective confidence translated into cautious responding. And third, when we compared the calibration curves for the slowest and fastest autistic individuals, the former curve was characterized by less overconfidence, a pattern that is inconsistent with an interpretation that their slowness reflected greater task difficulty. The convergence between our findings and those of Huggins and colleagues begs the question as to whether there might be considerable shared variance between measures of prospective confidence and self-report alexithymia measures that contain many items probing individuals about their perceptions of their own social–emotional competencies. Moreover, although global measures of metacognitive awareness may correlate negatively with measures of alexithymia (e.g., Babaei et al., 2016), they will not necessarily predict explicit measures of metacognitive awareness (such as used in our study), as has been highlighted by Kelly and Metcalfe (2011). In other words, investigations of the contributions of alexithymia to metacognitive awareness in processing emotion stimuli should incorporate specific tests of individuals' awareness of their processing, such as those provided by *G* coefficients and calibration approaches.

CONCLUSION

Our examination of metacognitive awareness of facial emotion recognition using different response formats, stimulus presentation types and an array of different emotions, provided no evidence of autism-related deficits. Autistic individuals were no poorer at discriminating correct from incorrect recognition responses than non-autistic individuals, although both groups exhibited marked inter-individual variability in discrimination performance. Further, confidence-accuracy calibration analyses revealed no autism–specific deficit in the sensitivity of monitoring of fluctuations in emotion recognition performance. The latter analyses also provided preliminary evidence that the slower recognition responding that characterized autistic individuals may reflect a difference of a strategic nature (e.g., greater caution) rather than a processing speed limitation.

ETHICS STATEMENT

The study was approved by the Social and Behavioral Ethics Committee of Flinders University.

ACKNOWLEDGMENTS

This research was supported by ARC DP190100162 and the Hamish Ramsay Fund. We thank Emily Barnett and Tess Barich for assistance with data and reliability coding and the participants from the autism research database who have assisted with our research program.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in Open Science Framework at https://osf.io/ndbfs/.

ORCID

Neil Brewer https://orcid.org/0000-0003-3465-9244
Carmen A. Lucas https://orcid.org/0000-0002-7984-8068
Marie Antonia Georgopoulos https://orcid.org/0000-0001-6744-3963
Robyn L. Young https://orcid.org/0000-0002-6124-8515

REFERENCES

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed.). American Psychiatric Publishing.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). American Psychiatric Publishing.

Babaei, S., Varandi, S. R., Hatami, Z., & Gharechahi, M. (2016). Metacognition beliefs and general health in predicting alexithymia in students. *Global Journal of Health Science*, *8*(2), 117–125. https://doi.org/10.5539/gjhs.v8n2p117

Bagby, R. M., Parker, J. D. A., & Taylor, G. J. (1994). The twenty-item Toronto alexithymia scale – I. Item selection and cross-validation of the factor structure. *Journal of Psychosomatic Research*, *38*(1), 23–32. https://doi.org/10.1016/0022-3999(94)90005-1

Baranski, J. V., & Petrusic, W. M. (1994). The calibration and resolution of confidence in perceptual judgments. *Perception & Psychophysics*, *55*(4), 412–428. https://doi.org/10.3758/BF03205299

Baranski, J. V., & Petrusic, W. M. (1995). On the calibration of knowledge and perception. *Canadian Journal of Experimental Psychology*, *49*(3), 397–407. https://doi.org/10.1037/1196-1961.49.3.397

Baranski, J. V., & Petrusic, W. M. (1998). Probing the locus of confidence judgments: Experiments on the time to determine confidence. *Journal of Experimental Psychology: Human Perception & Performance*, *24*(3), 929–945. https://doi.org/10.1037/–0096-1523.24.3.929

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

Brewer, N., & Smith, G. A. (1990). Processing speed and mental retardation: Deadline procedures indicate fixed and adjustable limitations. *Memory and Cognition*, *18*(5), 443–450. https://doi.org/10.3758/BF03198477

Brewer, N., & Wells, G. L. (2006). The confidence-accuracy relationship in eyewitness identification: Effects of lineup instructions, foil similarity and target-absent base rates. *Journal of Experimental Psychology: Applied*, *12*(1), 11–30.

Brewer, N., Young, R. L., Norris, J. E., Maras, K., Michael, Z., & Barnett, E. (2022). A quick measure of theory of mind in autistic adults: Decision accuracy, latency and self-awareness. *Journal of Autism and Developmental Disorders*, *52*(6), 2479–2496. https://doi.org/10.1007/s10803-021-05166-7

Brosnan, M., Ashwin, C., & Lewton, M. (2017). Brief report: Intuitive and reflective reasoning in autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *47*(8), 2595–2601. https://doi.org/10.1007/s10803-017-3131-3

Brosnan, M., Johnson, H., Grawemeyer, B., Chapman, E., Antoniadou, K., & Hollinworth, M. (2016). Deficits in metacognitive monitoring in mathematics assessments in learners with autism spectrum disorder. *Autism*, *20*(4), 463–472. https://doi.org/10.1177/1362361315589477

Brosnan, M., Lewton, M., & Ashwin, C. (2016). Reasoning on the autism spectrum: A dual process theory account. *Journal of Autism and Developmental Disorders*, *46*(6), 2115–2125. https://doi.org/10.1007/s10803-016-2742-4

Carpenter, K. L., Williams, D. M., & Nicholson, T. (2019). Putting your money where your mouth is: Examining metacognition in ASD using post-decision wagering. *Journal of Autism and Developmental Disorders*, *49*(10), 4268–4279. https://doi.org/10.1007/s10803-019-04118-6

Cook, R., Brewer, R., Shah, P., & Bird, G. (2013). Alexithymia, not autism, predicts poor recognition of emotional facial expressions. *Psychological Science*, *24*(5), 723–732. https://doi.org/10.1177/0956797612463582

Cooke, W. E. (1906). Forecasts and verification in Western Australia. *Monthly Weather Review*, *34*, 23–24.

Frith, U., & Happé, F. (1994). Autism: Beyond "theory of mind.". *Cognition*, *50*(1–3), 115–132. https://doi.org/10.1016/0010-0277(94)90024-8

Georgopoulos, M. A., Brewer, N., Lucas, C. A., & Young, R. L. (2022). Speed and accuracy of emotion recognition in autistic adults: The role of stimulus type, response format and emotion. *Autism Research*. https://doi.org/10.1002/aur.2713

Gigerenzer, G., Hoffrage, U., & Kleinbolting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review*, *98*(4), 506–528. https://doi.org/10.1037/0033-295X.98.4.506

Grainger, C., Williams, D. M., & Lind, S. E. (2014a). Metacognition, metamemory, and mindreading in high-functioning adults with autism Spectrum disorder. *Journal of Abnormal Psychology*, *123*(3), 650–659. https://doi.org/10.1037/a0036531

Grainger, C., Williams, D. M., & Lind, S. E. (2014b). Online action monitoring and memory for self-performed actions in autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *44*(5), 1–14. https://doi.org/10.1007/s10803-013-1987-4

Grainger, C., Williams, D. M., & Lind, S. E. (2016). Metacognitive monitoring and control processes in children with autism spectrum disorder: Diminished judgement of confidence accuracy. *Consciousness and Cognition*, *42*, 65–74. https://doi.org/10.1016/J.CONCOG.2016.03.003

Harms, M. B., Martin, A., & Wallace, G. L. (2010). Facial emotion recognition in autism spectrum disorders: A review of behavioral and neuroimaging studies. *Neuropsychology Review*, *20*(3), 290–322. https://doi.org/10.1007/s11065-010-9138-6

Huggins, C. F., Cameron, I. M., & Williams, J. H. G. (2021). Autistic traits predict underestimation of emotional abilities. *Journal of Experimental Psychology: General*, *150*(5), 930–942. https://doi.org/10.1037/xge0000784

Huggins, C. F., Donnan, G., Cameron, I. M., & Williams, J. H. G. (2021). Emotional self-awareness in autism: A meta-analysis of group differences and developmental effects. *Autism*, *25*(2), 307–321. https://doi.org/10.1177/1362361320964306

Juslin, P., Winman, A., & Olsson, H. (2000). Naive empiricism and dogmatism in confidence research: A critical examination of the hard-easy effect. *Psychological Review*, *107*(2), 384–396. https://doi.org/10.1037/0033-295X.107.2.384

Keating, C. T., Fraser, D. S., Sowden, S., & Cook, J. L. (2022). Differences between autistic and non-autistic adults in the recognition of anger from facial motion remain after controlling for alexithymia. *Journal of Autism and Developmental Disorders*, *52*(4), 1855–1871. https://doi.org/10.1007/s10803-021-05083-9

Kelly, K. J., & Metcalfe, J. (2011). Metacognition of emotional face recognition. *Emotion*, *11*(4), 896–906. https://doi.org/10.1037/a0023746

Koriat, A., & Ackerman, R. (2010). Choice latency as a cue for children's subjective confidence in the correctness of their answers. *Developmental Science*, *13*(3), 441–453. https://doi.org/10.1111/j.1467-7687.2009.00907.x

Lichtenstein, S., & Fischhoff, B. (1977). Do those who know more also know more about how much they know? *Organizational Behavior and Human Performance*, *20*(2), 159–183. https://doi.org/10.1016/0030-5073(77)90001-0

Lichtenstein, S., Fischhoff, B., & Phillips, L. D. (1982). Calibration of probabilities: The state of the art to 1980. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 306–334). Cambridge University Press.

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), 933–945. https://doi.org/10.1017/S0954579414000479

Mair, P., & Wilcox, R. (2019). Robust statistical methods in R using the WRS2 package. *Behavior Research Methods*, *52*(2), 1–25. https://doi.org/10.3758/s13428-019-01246-w

Maras, K., Gamble, T., & Brosnan, M. (2019). Supporting metacognitive monitoring in mathematics learning for young people with autism spectrum disorder: A classroom-based study. *Autism*, *23*(1), 60–70. https://doi.org/10.1177/1362361317722028

Maras, K., Norris, J. E., & Brewer, N. (2020). Metacognitive monitoring and control of eyewitness memory reports in autism. *Autism Research*, *13*(11), 2017–2029. https://doi.org/10.1002/aur.2278

McMahon, C. M., Henderson, H. A., Newell, L., Jaime, M., & Mundy, P. (2016). Metacognitive awareness of facial affect in higher-functioning children and adolescents with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *46*(3), 882–898. https://doi.org/10.1007/s10803-015-2630-3

Mottron, L., Dawson, M., Soulières, I., Hubert, B., & Burack, J. (2006). Enhanced perceptual functioning in autism: An update, and eight principles of autistic perception. *Journal of Autism and Developmental Disorders*, *36*(1), 27–43. https://doi.org/10.1007/s10803-005-0040-7

Nuske, H., Vivanti, G., & Dissanayake, C. (2013). Are emotion impairments unique to, universal, or specific in autism spectrum disorder? A comprehensive review. *Cognition & Emotion*, *27*(6), 1042–1061. https://doi.org/10.1080/02699931.2012.762900

O'Reilly, H., Lundqvist, D., Pigat, D., Baron, K., Fridenson, S., Tal, S., Meir, N., Berggren, S., Lassalle, A., Golan, O., Bolte, S., Piana, S., Rotman, C., Coletta, P., Marchi, E., Davies, I., Sullings, N., Baranger, A., Gauvain, C., … Baron-Cohen, S. (2012). *The EU-emotion stimulus set*. Autism Research Centre, University of Cambridge.

O'Reilly, H., Pigat, D., Fridenson, S., Berggren, S., Tal, S., Golan, O., Bölte, S., Baron-Cohen, S., & Lundqvist, D. (2016). The EU-emotion stimulus set: A validation study. *Behavior Research Methods*, *48*(2), 567–576. https://doi.org/10.3758/s13428-015-0601-4

Pachella, R. G. (1974). The interpretation of reaction time in information processing research. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition* (pp. 41–82). Erlbaum.

Palmer, M. A., Brewer, N., Weber, N., & Nagesh, A. (2013). The confidence-accuracy relationship for eyewitness identification decisions: Effects of exposure duration, retention interval, and divided attention. *Journal of Experimental Psychology: Applied*, *19*(1), 55–71. https://doi.org/10.1037/a0031602

Sauer, J., Brewer, N., Zweck, T., & Weber, N. (2010). The effect of retention interval on the confidence-accuracy relationship for eyewitness identification. *Law and Human Behavior*, *34*(4), 337–347. https://doi.org/10.1007/s10979-009-9192-x

Sawyer, A. C. P., Williamson, P., & Young, R. (2014). Metacognitive processes in emotion recognition: Are they different in adults with Asperger's disorder? *Journal of Autism and Developmental Disorders*, *44*(6), 1373–1382. https://doi.org/10.1007/s10803-013-1999-0

Semmler, C., Brewer, N., & Wells, G. L. (2004). Effects of postidentification feedback on eyewitness identification and nonidentification confidence. *Journal of Applied Psychology*, *89*(2), 334–346. https://doi.org/10.1037/0021-9010.89.2.334

Taylor, E. C., Farmer, G. D., Livingston, L. A., Callan, M. J., & Shah, P. (2022). Rethinking fast and slow processing in autism. *Journal of Psychopathology and Clinical Science*, *131*, 392–406. https://doi.org/10.1037/abn0000734

Tracy, J. L., Robins, R. W., Schriber, R. A., & Solomon, M. (2011). Is emotion recognition impaired in individuals with autism spectrum disorders? *Journal of Autism and Developmental Disorders*, *41*(1), 1517–1526. https://doi.org/10.1007/s10803-010-1030-y

Uljarević, 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

Weber, N., & Brewer, N. (2004). Confidence-accuracy calibration in absolute and relative face recognition judgments. *Journal of Experimental Psychology: Applied*, *10*(3), 156–172. https://doi.org/10.1037/1076-898X.10.3.156

Wechsler, D. (2011). *Wechsler abbreviated scale of intelligence–second edition (WASI-II)*. Pearson.

Williams, D. M., Bergström, Z., & Grainger, C. (2018). Metacognitive monitoring and the hypercorrection effect in autism and the general population: Relation to autism(−like) traits and mindreading. *Autism*, *22*(3), 259–270. https://doi.org/10.1177/1362361316680178

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Brewer, N., Lucas, C. A., Georgopoulos, M. A., & Young, R. L. (2022). Facing up to others' emotions: No evidence of autism-related deficits in metacognitive awareness of emotion recognition. *Autism Research*, *15*(8), 1508–1521. https://doi.org/10.1002/aur.2781