



Universiteit
Leiden

The Netherlands

Resonating emotions: an embodied perspective on alterations in facial emotion processing in autism and social anxiety

Folz, J.

Citation

Folz, J. (2025, January 17). *Resonating emotions: an embodied perspective on alterations in facial emotion processing in autism and social anxiety*.

Retrieved from <https://hdl.handle.net/1887/4177240>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4177240>

Note: To cite this publication please use the final published version (if applicable).

Chapter 5

**Individual differences in
interoception and autistic traits
share altered facial emotion
perception, but not recognition
per se**

Abstract

While alterations in both physiological responses to others' emotions as well as interoceptive abilities have been identified in autism, their relevance in altered emotion recognition is largely unknown. We here examined the role of interoceptive ability, facial mimicry, and autistic traits in facial emotion processing in non-autistic individuals. In an online Experiment 1, participants ($N = 99$) performed a facial emotion recognition task, including ratings of perceived emotional intensity and confidence in emotion recognition, and reported on trait interoceptive accuracy, interoceptive sensibility and autistic traits. In a follow-up lab Experiment 2 involving 100 participants, we replicated the online experiment and additionally investigated the relationship between facial mimicry (measured through electromyography), cardiac interoceptive accuracy (evaluated using a heartbeat discrimination task), and autistic traits in relation to emotion processing. Across experiments, neither interoception measures nor facial mimicry accounted for a reduced recognition of specific expressions with higher autistic traits. Higher trait interoceptive accuracy was rather associated with more confidence in correct recognition of some expressions, as well as with higher ratings of their perceived emotional intensity. Exploratory analyses indicated that those higher intensity ratings might result from a stronger integration of instant facial muscle activations, which seem to be less integrated in intensity ratings with higher autistic traits. Future studies should test whether facial muscle activity, and physiological signals in general, are correspondingly less predictive of perceiving emotionality in others in individuals on the autism spectrum, and whether training interoceptive abilities might facilitate the interpretation of emotional expressions.

Based on: Folz, J., Nikolić, M., & Kret, M. E. (2024). Individual differences in interoception and autistic traits share altered facial emotion perception, but not recognition per se. Manuscript conditionally accepted for publication in Scientific Reports

Data availability statement:

The datasets and materials generated and/or analysed during the current study will be made available on Dataverse NL upon publication: <https://doi.org/10.34894/TOUSTD>

Supplementary material:



Difficulties in recognizing others' emotions have been assumed to be one relevant source of broader social interaction difficulties in individuals on the autism spectrum (D. A. Trevisan & Birmingham, 2016; choice of autism terminology informed by Botha et al., 2023). Yet, recent research suggests a more differentiated picture: different paths to emotion recognition - rather than differences in ability- are likely to explain observed differences in emotion recognition between individuals on the autism spectrum and non-autistic individuals (Arnaud, 2020; Stel et al., 2008). For example, individuals on the autism spectrum might integrate their own mental representations of emotions less (Keating et al., 2023) and rather rely on learned rules when interpreting emotional expressions (Rutherford & McIntosh, 2007). One other path to emotion recognition draws on interoception, which involves the sensation of (changes in) physiological states (Craig, 2002). Via automatic alignment to an expressed emotion, or so-called emotional contagion (E. Hatfield et al., 1993; Prochazkova & Kret, 2017), physiological changes can not only inform an individual about their own emotional experience (Damasio, 1996) but can also offer insights in the emotional experience of others via simulation, such as facial mimicry (Wood et al., 2016). While previous research suggests that physiological responses to others' emotions (Davies et al., 2016; Hubert et al., 2009), as well as interoception (Garfinkel et al., 2016), would be altered in autism, little is known about the relevance of these alterations in an emotion recognition context. In the current study with a non-autistic sample, we aimed to approach a better understanding of the role of physiological signals and their sensation in emotion recognition alterations in relation to autistic trait levels paving the way for future investigations in autism..

The Body in Emotion Perception in Autism

Past studies on the recognition of facial or bodily emotion expressions in autism predominantly report worse performance, that is, a lower sensitivity to emotions or less accuracy in labelling them, compared to non-autistic samples (Frank et al., 2018; Uljarevic & Hamilton, 2013). Individuals on the autism spectrum have further been shown to differ from non-autistic individuals in their physiological responses to observed emotional expressions. More specifically, both hyper- and hypoarousal to emotion displays have been reported (Cuve et al., 2018; Hubert et al., 2009), whereas the automatic mirroring of facial expressions (i.e., facial mimicry) has typically been found to be reduced (Davies et al., 2016). Facial mimicry patterns are thought to play an important role in attributing discrete emotions to

facial expressions, as they are specific for different emotion categories (Folz et al., 2022; Wingenbach et al., 2020) and can act as simulations of observed expressions (Arnold & Winkelman, 2020; Wood et al., 2016). Robust evidence for a link between facial mimicry and emotion recognition has not been established in the scarce literature on this topic (Holland et al., 2020). Yet, some studies have found that non-autistic individuals were influenced by their own facial expressions in attributing emotional states to themselves (Soussignan, 2002; Stel & van Knippenberg, 2008) and others (Drimalla et al., 2019; Sato et al., 2013). In contrast, no influence of facial mimicry on experienced emotions, even if it was intentionally produced, has been reported in individuals on the autism spectrum (Stel & van Knippenberg, 2008). Furthermore, in non-autistic samples, reduced facial mimicry has, if at all, only been linked to higher autistic trait levels for very specific emotions and subgroups (Hermans et al., 2009). Interestingly, in our recent study in which no systematic modulations of facial mimicry by autistic trait levels have been found, a weaker link between facial mimicry responses to sad facial expressions (i.e., mirroring of frowns) and successful emotion recognition has been observed (Folz et al., 2023). Thus, the presence of physiological alignment to emotional expressions might not be sufficient to facilitate emotion recognition. In order to integrate information about one's own physiological state in emotion processing, certain interoceptive abilities, namely an awareness of changes in physiological state as well as an accurate representation thereof, may be necessary. Findings from various studies support the link between emotional and somatic awareness, with the latter being more fundamental (Kanbara & Fukunaga, 2016). Hence, in the current study we aimed to identify whether individual differences in the sensation and integration of one's physiological signals would be linked to emotion recognition outcomes, and whether this could offer an explanation to altered emotion processing associated with variations in autistic trait levels.

Interoception in Emotion Processing

Research on interoception, or the "sense of the physiological condition of the entire body" (Craig, 2002), has recently highlighted the integration of physiological signals in central processing beyond homeostatic control, widely influencing human cognition and behaviour (Critchley & Harrison, 2013). This also entails the affective domain (Critchley & Garfinkel, 2017). By detecting and assigning meaning to physiological changes, interoceptive processes can become an important mechanism in emotion processing (Smith & Lane, 2015). From a predictive coding

perspective, emotional states have even been suggested to arise from active inference of causes of physiological changes (Seth, 2013). If, for example, various afferent interoceptive signals indicate a state of heightened physiological arousal, the mismatch to a predicted calmer state is resolved by acquiring more sensory information about likely internal and external causes, with their integration in updated models resulting in an emotion percept. Importantly, individuals vary in the processing of interoceptive input at different levels in the hierarchy (Suksasilp & Garfinkel, 2022), and various measures have been developed to assess these individual differences. In a common interoception model (Garfinkel et al., 2016), three interoceptive dimensions are distinguished (Forkmann et al., 2016; Garfinkel et al., 2015), namely interoceptive accuracy (i.e., the objective accuracy in the detection of interoceptive signals), interoceptive sensibility (i.e., the self-reported, subjective tendency to focus and be aware of interoceptive signals) and interoceptive awareness (i.e., the ability to assess one's interoceptive accuracy correctly, i.e. a metacognitive process). Next to these dimensions, the correspondence between interoceptive sensibility and accuracy, the so-called interoception trait prediction error (Garfinkel et al., 2016), can provide valuable information about the mismatch between subjective beliefs and objective measures. The scope of this model in describing subjective beliefs is limited as it fails to distinguish between beliefs regarding accuracy in perceiving interoceptive signals versus attention to them. To capture this dissociation, Murphy et al. (2019) developed a 2x2 factor model of interoceptive ability, with the first factors ('What is measured?') distinguishing between accuracy and attention. The second factor ('How is it measured?') contrasts beliefs regarding one's accuracy/attention (i.e., self-reports) with one's actual behaviour regarding the two targets (i.e., objective measures).

Most research in the field of emotion processing has employed objective task-based interoception measures, which contrast the (objectively measured) nature of a specific afferent signal (e.g., timing, strength) to its subjective experience. The firing of baroreceptors has been highlighted as afferent signal in the cardiovascular domain, among other signals (Desmedt et al., 2023), indicating cardiovascular arousal in emotion processing (Critchley & Garfinkel, 2017). In frequently used cardiac interoception tasks, participants either keep track of their heartbeats within a specific time window (i.e., heartbeat counting) or judge the synchronicity of their heartbeats with auditory information (i.e., heartbeat discrimination). Even

though these tasks were designed to provide insights in the subjective experience of afferent cardiac signals, they are not exclusively reflecting the accurate perception of cardiac interoceptive signals. The heartbeat discrimination task, for example, further requires participants to match the subjective experience of the heartbeat timing to the timing of an external stimulus. Diverging task demands might thus also explain the relatively low correspondence between different measures of cardiac interoceptive accuracy (Hickman et al., 2020). Interoceptive accuracy can be assessed in various bodily systems (i.e., domains; Murphy et al., 2018), and performance on objective interoceptive accuracy measures in these distinct domains (e.g., cardiac, respiratory) differs within individuals (Garfinkel et al., 2016). Self-report measures on interoceptive accuracy, as described in the 2x2 factor model, aim to assess the accurate perception of interoceptive signals across domains (referred to as “trait interoceptive accuracy” in the following; Murphy et al., 2019). The Interoceptive Accuracy Scale requires participants to rate the degree to which their subjective experience of several afferent signals each relate to actual physiological needs. Indicators for evaluating the accurate perception of sensations can be quite diverse, with actually vomiting (when feeling the urge to vomit) or being full (after giving in to hunger) as examples. Although this multifaceted measure of interoceptive accuracy may capture not only interoception but also subjective beliefs and experiences in everyday life, there is evidence that it would correspond to cardiac interoceptive accuracy (Murphy et al., 2020).

In previous research on the role of interoception in emotion processing, individuals with higher cardiac interoceptive accuracy have not only been found to show stronger physiological responses (Pollatos & Schandry, 2008) and report more intense emotional experiences (Dirupo et al., 2020; Wiens et al., 2000) when viewing emotional images, but the link between their physiological changes and their subjective arousal levels has also been reported to be stronger (Dunn et al., 2010). This is in line with the suggestion that individuals with high objective interoceptive accuracy would be able to increase the precision of their interoceptive prediction errors relative to their interoceptive priors, and also to other sensory modalities, via attentional processes (Ainley et al., 2016): once physiological changes are detected and propagated in an emotional context, individuals with higher objective interoceptive accuracy should show stronger autonomic responses to emotional stimuli, (i.e., a reinforcement via active inference) as bottom-up interoceptive information should have a stronger influence on

information processing. Higher cardiac interoceptive accuracy has further been related to a better recognition of negative emotional expressions (Fittipaldi et al., 2020; Terasawa et al., 2014), supporting the idea that an accurate representation of interoceptive information might also facilitate recognizing emotional states of others. The few studies examining the role of interoceptive sensibility in emotion processing have also mainly observed a facilitation of processing, such as faster emotion recognition (Hübner et al., 2021) and a more precise adaption to emotion probabilities (Hübner et al., 2021, 2022). Yet, links might be specific to different subcomponents of interoceptive sensibility (Desdentado et al., 2022) as well as depend on the task at hand (Ventura-Bort et al., 2021). Overall, individuals with a less accurate representation of interoceptive signals or a lower tendency to monitor them might not benefit from their integration in emotion recognition, as it might be the case in autism.

Interoception in Autism

Alterations in interoception have been associated with various physical, neurodevelopmental and mental health conditions (Bonaz et al., 2021; Khalsa et al., 2018), including autism (DuBois et al., 2016; Proff et al., 2022). Compared to non-autistic control samples, many studies have found a reduced interoceptive accuracy in adults (Failla et al., 2020; Garfinkel et al., 2016; Mul et al., 2018) and children (Failla et al., 2020; Nicholson et al., 2019) on the autism spectrum. Worse performance in interoceptive accuracy tasks could, however, not consistently be observed in both populations (Nicholson et al., 2019; Schauder et al., 2015), and also when using different tasks (Z. J. Williams et al., 2023). Studies examining the subjective experience of interoceptive signals (i.e., sensibility) are similarly inconsistent: studies have found increased sensibility (Garfinkel et al., 2016; Pickard et al., 2020), reduced sensibility (Mul et al., 2018) or no differences between individuals on the autism spectrum and non-autistic individuals (Butera et al., 2023). Different study populations as well as measurement tools might explain inconsistencies. Questionnaires focusing on the sensation of specific body signals, such as the Body Perception Questionnaire (Porges, 1993), might be reflective of the hypersensitivity to interoceptive signals that individuals on the autism spectrum can experience. This increased interoceptive sensibility has been found to strongly diverge from a decreased interoceptive accuracy in individuals on the autism spectrum, resulting in a relatively higher interoceptive trait prediction error (Garfinkel et al., 2016). In contrast, questionnaires focusing on a more global

awareness, integration or interpretation of signals, such as the Multidimensional Assessment of Interoceptive Awareness (Mehling et al., 2018), might rather be reflective of the difficulty to make sense of bodily signals. In order to capture these subjectively experienced difficulties in individuals on the autism spectrum, the Interoception Sensory Questionnaire was developed as assessment tool for “interoceptive confusion”. Here, interoceptive confusion has not only been found to be highly prevalent in individuals on the autism spectrum, but also increasing in severity the higher an individual’s autistic trait levels in a non-autistic population (Fiene et al., 2018). These findings are in line with predictive coding theories on interoception in autism (Quattrocki & Friston, 2014; van de Cruys et al., 2014) which suggest that individuals on the autism spectrum would be hypersensitive to interoceptive signals, overrepresent them at low processing levels (i.e., distinct sensations) and have a reduced accuracy in their sensation due to highly precise and inflexible prediction errors, while the integration of signals to a global awareness might be constrained. As previously outlined, somatic awareness might build an important foundation for emotional awareness with regard to both our own and others’ emotions (Kanbara & Fukunaga, 2016). Thus, sensing interoceptive signals less accurately or integrating them to a lesser degree could potentially explain differences in emotion recognition tasks that are observed between non-autistic individuals and individuals on the autism spectrum, or related to high autistic trait levels.

The Role of Interoception in Processing Others’ Emotions in Autism

Only few studies have investigated the role of interoception in processing others’ emotions in autism. Focusing on (emotional) empathy as an outcome, two recent studies comparing individuals on the autism spectrum to non-autistic individuals have shown inconsistent findings: While no group differences in interoceptive sensibility measures, as well as no link to emotional empathy, have been observed in one study (Butera et al., 2023), the other study has found both reduced interoceptive sensibility and cardiac interoceptive accuracy in autism, with the latter showing a negative relation with empathy (Mul et al., 2018). Importantly, both studies have highlighted the relevance of co-occurring alexithymia, a trait encompassing difficulties in identifying and describing one’s emotions (Nemiah et al., 1976), in explaining the link between altered interoceptive processing and potential difficulties in empathy related to autism. Alexithymia has consistently been linked to difficulties in emotion recognition (Jongen et al., 2014; Lane

et al., 1996; Parker et al., 1993) and shows a high prevalence in autism (49.93%; Kinnaird et al., 2019). Studies assessing alexithymia levels in individuals on the autism spectrum provide evidence that alterations in various aspects of emotion processing in autism (Gaigg et al., 2018; Ketelaars et al., 2016), as well as in interoceptive ability (Shah et al., 2016), might indeed be explained by co-occurring high alexithymia levels. Yet, whether a reduced subjective and/or objective interoceptive accuracy would account for difficulties in emotion recognition related to autism remains an open question. In contrast to this potential consequences of a reduced interoceptive accuracy, The heightened interoceptive sensibility in autism, reflecting a hypersensitivity to (specific) bodily signals, has been linked to more severe autism symptomology in specific domains, namely to socio-affective features in children (Palser et al., 2020) as well as to a reduced emotion sensitivity and the occurrence of anxiety symptoms in adults on the autism spectrum (Garfinkel et al., 2016). Thus, learning to regulate and optimally integrate interoceptive information might benefit individuals on the autism spectrum in their daily (social) functioning and experiences. While the amount of literature on clinical interventions in autism focusing on attention to and integration of physiological signals is growing (Gaigg et al., 2020; Quadt et al., 2021), the role of altered interoceptive processing in autism symptomology, including the socio-affective domain, is still scarcely investigated.

Individual characteristics associated with autism can be observed in the general population to varying degrees, resulting in claims that individuals on the autism spectrum could be positioned at the extreme of a continuum of autistic traits (Constantino & Todd, 2003; Robinson et al., 2011). This perspective received support by genetic studies (Bralten et al., 2018; Lundström et al., 2012) as well as studies focusing on behavioural aspects of autism (Mayer, 2017). Accordingly, non-autistic individuals with higher autistic trait levels show, in some regards and to some extent, similar patterns of alterations in processing observed emotional expressions as individuals on the autism spectrum (Åsberg Johnels et al., 2017; Folz et al., 2023; Hermans et al., 2009). Importantly, findings on links between autistic trait levels in non-autistic samples and certain outcomes of interest cannot simply be generalized to autism, let alone to experiences of individuals on the autism spectrum (Sasson & Bottema-Beutel, 2022). They can, however, help forming assumptions on which factors, within processes that show similar

patterns of alterations in autism and high autistic trait levels, might be relevant to further examine in autism (Pollmann et al., 2010).

Objectives of the Current Study

We investigated the role of interoception and facial mimicry in emotion processing in relation to autistic trait levels in two pre-registered experiments with non-autistic individuals (see Figure 1): The first, online experiment consisted of a facial emotion recognition task with confidence judgments in the accuracy of recognition and intensity ratings of seen expressions as well as questionnaires on autistic traits, trait interoceptive accuracy and interoceptive sensibility. Social anxiety traits and alexithymia were assessed as control variables given that both have been related to alterations in interoception (Desai et al., 2019; Stevens et al., 2011), as well as difficulties in the socio-affective domain in autism (Mul et al., 2018; Spain et al., 2018). In our main analysis, we examined whether a reduced trait interoceptive accuracy would explain a reduced emotion recognition accuracy with higher autistic trait levels, while controlling for alexithymia and social anxiety traits. In the second, lab-based experiment, we expanded the online study. Next to assessing the same measures as in the online experiment for replication purposes, we added facial electromyography recordings during the emotion recognition task. This allowed us to investigate a second factor which could play a role in reduced emotion recognition with higher autistic trait levels: We tested whether physiological responses to others' facial expressions (i.e., facial mimicry) were less predictive of emotion recognition accuracy with higher autistic trait levels as found in a previous study for sad facial expressions (Folz et al., 2023). By adding a heartbeat discrimination task to the lab experiment, we could further explore whether autistic traits would be linked to lower cardiac interoceptive accuracy and/or a stronger mismatch between subjective and objective measures of interoception (interoceptive trait prediction error).

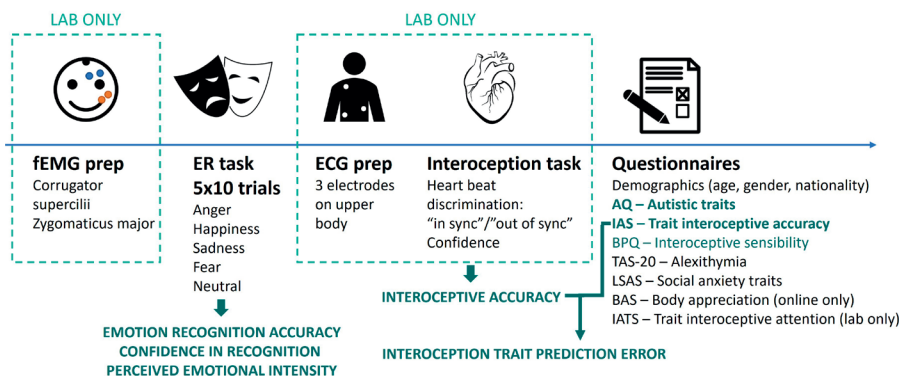


Figure 1. Experimental design of the online study (Experiment 1) and the lab study (Experiment 2. prep = preparation, fEMG = facial electromyography, ER = emotion recognition, ECG = electrocardiography, AQ = Autism Quotient, IAS = Interoceptive Accuracy Scale, BPQ = Body Perception Questionnaire, LSAS = Liebowitz Social Anxiety Scale, TAS-20 = Toronto Alexithymia Scale (20-item version), BAS = Body Appreciation Scale (not examined), IATS = Interoceptive Attention Scale.

Experiment 1

Method

Participants

We tested 100 adult participants between 18-35 years-old who reported no prior or current psychiatric or neurological disorder. The choice of the sample size was based on a power analysis with simulated data, indicating a power of .94 (or 1) to find significant relations between autistic traits (or social anxiety traits) and emotion recognition accuracy with a similar effect size as in a previous study (Folz et al., 2023). More details on the sample size rationale can be found in the preregistration of the study (<https://osf.io/wugq7>). Out of the 100 participants, 70 participants were recruited via the online recruitment platform SONA of Leiden university (student population) and 30 participants were recruited via a direct link between 28/12/2020 and 24/01/2021. One participant did not meet the age criterion (18-35 years-old) and was excluded after data collection. Our final sample consisted of 99 participants (86 females, 12 males, 1 'prefer not to say') with an age range between 18 and 34 years ($M = 21.39$, $SD = 4.27$). The majority of our participants were Dutch (45 individuals), Macedonian (23 individuals) or German (10 individuals), and all participants completed the experiment in English. There

was no direct monetary reimbursement for participation. Yet, all participants could enter a lottery (10% chance of winning) to either receive a 10€ vouchers for an online store in the EU or to donate 10€ to 'Give Well' charity (<https://www.givewell.org/>). Leiden university students could additionally receive 2 course credits. All participants provided informed consent prior to participation. The study was conducted according to the Declaration of Helsinki and the protocol was approval by the local ethics committee of the Faculty of Social and Behavioural Sciences at Leiden University (2020-12-18-M.E. Kret-V1-2834).

Stimuli

Color videos of 5 male and 5 female individuals, showing spontaneous facial expressions of anger, happiness, fear, sadness and neutral in full frontal view, were selected from a previous standardization (Folz et al., 2023) of the FEEDTUM stimulus database (Wallhoff et al., 2006). All videos had a length of 2s (500ms neutral expression followed by 1500ms expression of the respective category) and a gray background. The size of the videos was automatically adjusted to the participants viewport size (4: 3 ratio). All participant viewed all 50 videos once in a random order.

Procedure

The experiment was performed on the online experiment platform Gorilla (Anwyl-Irvine et al., 2020) (<https://gorilla.sc/>). Participants were instructed to complete it on a PC screen in a quiet room, without any disturbances. As first part of the experiment, participants performed an emotion recognition task. Each trial started with a central fixation cross for 1s. Afterwards, a facial expression video (see Stimuli section) was presented in the centre of the screen for 2s and followed by a 100ms blank screen. On the next screen, participants chose a label for the displayed expression ('Which type of expression was displayed by the person in the video?') out of the 5 potential categories (angry/happy/sad/fearful/neutral). They additionally rated the confidence in their decision ('How confident are you about your decision?') on a visual analogue scale from 'not confident at all' to 'very confident'. Integer values ranged from 0-100 but were not visible to the participant. As last rating on the same screen, participant indicated the perceived emotional intensity of the expression ('How emotionally intense was the expression displayed in the video?') on a visual analogue scale from 'not intense at all' to 'very intense'. Integer values again ranged from 0-100 but were not visible to the participant. Participants could move on to the next trial in a self-paced manner once all three questions had been answered. After the emotion

recognition task, all participants provided demographical information regarding their age, gender, and nationality first. The order of the following questionnaires (see Measurements section and Figure 1) was randomized across participants. At the end of the experiment, participants could decide to enter the lottery for the 10€ vouchers/donation by providing their email address.

Measurements

Autistic traits. We used the Autism-Spectrum Quotient (AQ; Baron-Cohen et al., 2001) as a self-report measure of traits associated with the autism spectrum. Respondents rate how strongly each of 50 items applies to them on a 4-point Likert scale (1 = definitely agree, 2 = slightly agree, 3 = slightly disagree, and 4 = definitely disagree). Some items are reverse-coded, and all items scores are binarized (1 or 2 to 0 and 3 or 4 to 1) before summation. Sum scores can be calculated for five separate subscales with ten items each (social skill, attention switching, attention to detail, communication, and imagination) as well as for one total autistic trait score. Higher sum scores reflect higher autistic trait levels. In our experiments, 2 participants had a higher AQ score than 32 which has been described as cut-off for clinical significance. More detailed descriptive information about all questionnaires scores, including an overview of the reliabilities and distribution parameter, can be found in Table 1. A visualization of the relations between the questionnaire measures can be found in Figure S1A in the Supplemental Materials.

Trait interoceptive accuracy. The Interoceptive Accuracy Scale (IAS; Murphy et al., 2020) was used to assess self-reported interoceptive accuracy with regard to various body sensations (e.g., heartbeat, hunger, need to urinate,...). Interoceptive accuracy for each of the 21 IAS-items are evaluated on a 5-point Likert scale (1 = disagree strongly, 2 = disagree, 3 = neither agree nor disagree, 4 = agree, and 5 = strongly agree), and higher sum scores represent a higher self-reported interoceptive accuracy (see Table 1).

Interoceptive sensibility. We used the body awareness scale of the short form of the Body Perception Questionnaire (BPQ-SF; Porges, 1993) to assess interoceptive sensibility in our sample. Statements regarding the awareness of 26 body sensations (e.g., sweaty palms, stomach and gut pains,...) are rated on a 5-point Likert scale (1 = never, 2 = occasionally, 3 = sometimes, 4 = usually, and 5 = always). Sum scores are regarded as an integrated measure of interoceptive sensibility, with higher scores indicating higher interoceptive sensibility (see Table 1).

Alexithymia. With the 20 item version of the Toronto Alexithymia Scale (TAS-20; Bagby et al., 1994), we assessed alexithymia in our sample. Each item of the TAS-20 is rated on a 5-point Likert scale from 1 (strongly disagree) to 5 (strongly agree), with five items being reversely coded. The items can be summarized on three subscales (“Difficulty identifying feelings”, “Difficulty describing feelings”, “Externally-oriented thinking”) as well as summed to a total score, with a higher score higher alexithymic trait levels. Total scores were higher than 61 for 10 participants (10%), indicating alexithymia (see Table 1).

Social anxiety traits. The Liebowitz Social Anxiety Scale (LSAS; Liebowitz, 1987) was used to assess self-reported social anxiety traits. Respondents rate their fear and avoidance in 24 social interaction and performance separately on 4-point Likert scales (fear: 1 = None, 2 = Mild, 3 = Moderate, 4 = Severe; avoidance: 1 = Never, 2 = Occasionally, 3 = Often, 4 = Usually). In our sample, 65 participants (66 %) exceeded the theoretical cut-off of 30 which indicates a probability of social anxiety disorder, with 26 participants (26%) scoring high (above 60) on this scale (see Table 1).

Body appreciation. As part of a Master thesis project, participants also completed the updated version of the Body Appreciation Scale (BAS-2; Tylka & Wood-Barcalow, 2015). The results were not of interest to answer the research question of this article and are therefore not reported.

Table 1. Distribution of Questionnaires Scores, Including their Internal Consistency (Cronbach’s Alpha) and Distribution Parameters, in the Online Study (Experiment 1)

	N	Mean	SD	Min	Max	α	ICL	hCL	kurtosis	skewness
AQ	99	17.05	6.59	3	34	0.79	0.73	0.85	2.89	0.37
IAS	99	79.80	9.25	51	100	0.82	0.77	0.87	2.97	-0.20
BPQ	99	72.48	25.16	26	126	0.96	0.95	0.97	2.00	0.15
LSAS	99	44.00	23.88	0	117	0.96	0.94	0.97	3.09	0.71
TAS	99	45.22	11.28	24	73	0.85	0.81	0.89	2.23	0.34

Note. AQ = Autism Quotient (Autistic traits), IAS = Interoceptive Accuracy Scale (Trait interoceptive accuracy), BPQ = Body Perception Questionnaire (Interoceptive sensibility), LSAS = Liebowitz Social Anxiety Scale (Social anxiety traits), TAS = Toronto Alexithymia Scale (Alexithymia), ICL/hCL = lower/higher Confidence Level of 95% Confidence Interval associated with α .

Data Analysis

We preregistered the data analyses to test our hypotheses on the Open Science Framework (<https://osf.io/wuqg7>). The data of the two experiments was collected at different stages of the Covid-19 pandemic. Since this might have resulted in biased replies on the social anxiety trait measure, which included, for example, questions about avoidance of social situations, we decided to focus on *Autistic traits* as the main predictor in our analyses, which were all conducted in R 4.2.2 (R Core Team, 2022). As preregistered, interactions between *Social anxiety traits* and *Emotion category* were still included in all clinical-trait-score-related analyses as control predictors, similar to *Alexithymia*. The two clinical trait score measures showed to have significant medium positive correlations with one another, as well as with *Alexithymia* (LSAS-AQ: $r_s = 0.47, p < .001$; LSAS-TAS: $r_s = 0.34, p < .001$; AQ-TAS: $r_s = 0.30, p = .003$), supporting our approach to control for *Social anxiety traits* and *Alexithymia* in all models. Before fitting our models, all continuous variables were standardized (i.e., centered and scaled) to obtain standardized beta coefficients. In order to test whether trait interoceptive accuracy would (partially) mediate the link between autistic trait levels and emotion recognition accuracy, we fitted three models using the lmerTest package (Kuznetsova et al., 2017): First, we tested whether emotion recognition accuracy was decreased with higher autistic trait levels (path c) while controlling for social anxiety traits and alexithymia. Previous literature has reported emotion-specific alterations in recognition performance with regard to autistic traits, but also with regard to social anxiety traits. Therefore, the binary outcome *Emotion recognition accuracy* (1 = correct, 0 = incorrect) was predicted by a two-way interaction between *Emotion category* (angry, happy, fearful, sad and neutral) and *Autistic traits* as well as by a two-way interaction between *Emotion category* and *Social anxiety traits*, and *Alexithymia* as control predictors. Random intercepts for each stimulus (50 stimuli) as well as for each participant (99 participants) were added. After model fitting, slopes for the relation between *Autistic traits* and *Emotion recognition accuracy* were estimated for each level of *Emotion category*, using the emtrends function of the emmeans packages (Lenth, 2023; Holm method for p-value adjustment). Second, we examined whether *Trait interoceptive accuracy* was reduced with higher *Autistic traits* levels (path a), while controlling for *Social anxiety traits* and *Alexithymia*. A linear regression analysis was performed with *Trait interoceptive accuracy* as outcome variable and *Autistic traits* as predictor of interest, as well as *Alexithymia* and *Social anxiety traits* as control predictors. In the third and final model fit, we added an interaction between

Emotion category and *Trait interoceptive accuracy* as a mediator of the association between *Autistic Traits* and *Emotion Recognition*, next to the predictors in the first model, to be able to identify whether the effect of *Autistic traits* on *Emotion recognition accuracy* for certain levels of *Emotion Category* was mediated by *Trait interoceptive accuracy* (path ab). The causal mediation model was tested using the RMediation package (Tofighi, 2023). From the previously defined models, path a was defined as the effect of *Autistic traits* on *Trait interoceptive accuracy* and path b as the effect of *Trait interoceptive accuracy* on *Emotion recognition accuracy* of expression(s) that were less well recognized with higher autistic trait levels. The indirect effect (path ab) of *Trait interoceptive accuracy* on the association between *Autistic traits* and *Emotion recognition accuracy* of (certain) emotional expressions was also tested for significance.

We further explored the role of both autistic traits as well as self-reported interoception measures in determining the two other emotion recognition task outcomes, namely confidence in emotion recognition and perceived emotional intensity of seen expressions. As we did not expect the variables to influence each other in predicting the outcomes and aimed to avoid inflation of type I error, all predictor variables were included in one mixed model for each outcome. *Perceived emotional intensity* was thus predicted by two-way interactions between *Emotion category* and *Autistic traits*, *Emotion category* and *Trait interoceptive accuracy*, and *Emotion category* and *Interoceptive sensibility* as well as by the two-way interaction between *Emotion category* and *Social anxiety traits*, and *Alexithymia* as control predictors. In line with the *Emotion recognition accuracy* models, random intercepts for each stimulus and each participant were added. The distribution of the confidence data was not normal and the highest value (100) was selected in many trials (20%), indicating full confidence. Therefore, we fitted a Bayesian GLMM, using the brms package (Bürkner, 2017), with a zero-one-inflated family instead of a LMM to predict *Confidence in emotion recognition* with the same random and fixed effect structure as the intensity model. Thus, we estimated separate parameters for a beta regression excluding zeros and ones (ϕ), for the proportion of zeros and ones only (zoi) as well as for the proportion of ones versus zeros (coi). Integrated posterior estimates for the slopes at three different values of the predictors of interest (-1, 0, 1) were obtained using the emtrends function of the emmeans package. All visualizations of effects and all model fit tables are based on the sjPlot package (Lüdtke, 2021).

As additional analyses, we explored relations between *Autistic traits* and *Interoceptive sensibility* as well as between *Interoceptive sensibility* and *Emotion recognition accuracy*. As no significant effects related to autistic trait levels were observed and these analyses were exploratory, the models are reported in the Supplemental Materials (see Tables S10 - S12) and not discussed in detail here.

Results

Main Analysis

We did not find evidence for trait interoceptive accuracy mediating the effect of autistic traits on emotion recognition (see Figure 2). As only recognition of angry expressions showed to be worse with higher autistic trait levels, (**path c**, see Figure 4A and Table 2 for the slope comparisons, as well as Table S1 in the Supplemental Materials for the full model fit), we tested for an indirect effect of *Trait interoceptive accuracy* on the association between *Autistic Traits* and *Emotion recognition accuracy* of angry expressions (**path ab**). Confidence limits included zero, indicating no mediated effect (see Figure 2). Thus, autistic traits predicted worse recognition of angry faces only directly. Against our expectations, *Autistic traits* was not a significant predictor of *Trait interoceptive accuracy* in our second model (**path a**, see Table S3 in the Supplemental Material), and neither *Trait interoceptive accuracy* nor its interaction with *Emotion category* were significant in predicting *Emotion recognition accuracy* (**path b**, see Table S5 in the Supplemental Material). An exclusion of *Alexithymia* from all models including *Autistic traits* as predictor did not result in a meaningful change of the outcomes. All model fits, including significant effects that are unrelated to our predictors of interest, can be found in Tables S1-5 in the Supplemental Materials.

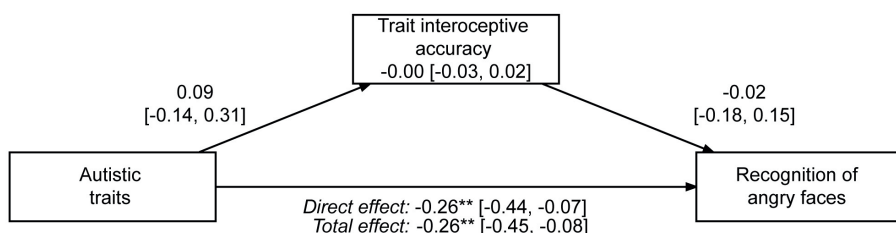


Figure 2. Results of the mediation analysis in Experiment 1 with trait interoceptive accuracy as potential mediator explaining worse recognition of angry faces with higher autistic trait levels. ** $p < .001$

Table 2. Estimated Slope of the Linear Relation Between Autistic Traits and Emotion Recognition by Emotion Category, and Results of the Slope Comparisons against Zero

	Estimated slope [SE]	95% CI	z-ratio	p-value
<i>Emotion category</i>				
Angry	-0.30 [0.11]	-0.58, -0.02	-2.78	.027
Fearful	0.02 [0.13]	-0.31, 0.36	0.18	1
Happy	-0.03 [0.21]	-0.57, 0.51	-0.16	1
Sad	-0.02 [0.10]	-0.28, 0.24	-0.18	1
Neutral	0.14 [0.11]	-0.15, 0.43	1.22	.895

Note. Confidence intervals are adjusted for multiple comparisons using the Bonferroni-method, and p-values are adjusted using the Holm-method.

Exploratory Analyses

In the model with *Perceived emotional intensity* as an outcome, we did not find an effect of *Autistic traits* (neither as main effect nor as interaction with *Emotion category*). While the interaction between *Trait interoceptive accuracy* and *Emotion category* was significant (see Figure 4C), $F(4,4786) = 3.28$, $p = .011$, the slope of the relations between *Trait interoceptive accuracy* and *Perceived emotional intensity* was not significantly different from 0 for any level of *Emotion category* (see Table S7 in the Supplemental Materials). Furthermore, we observed both a significant effect of the predictor *Interoceptive sensibility*, $F(1,93) = 4.01$, $p = .048$, as well as a significant interaction between *Emotion category* and *Interoceptive sensibility*, $F(4,4786) = 3.28$, $p = .011$, in predicting *Perceived emotional intensity*. The slope comparisons by *Emotion category* revealed that both neutral and sad expressions were perceived as more emotionally intense with higher *Interoceptive sensibility*: The slope for sad expressions was 0.14, 95% CL [0.00, 0.27], $t(145) = 2.65$, $p = .036$, and the slope for neutral expressions was 0.14, 95% CL [0.01, 0.28], $t(145) = 2.81$, $p = .027$ (see Figure 5A). Similar to the main analysis, excluding *Alexithymia* did not change the outcomes in a meaningful way. The complete model fits, including effects that are unrelated to our predictors of interest, can be found in Tables S6 and S8 in the Supplemental Materials.

In the Bayesian LMM with *Confidence in emotion recognition* as outcome, estimated slopes at three points of the *Autistic traits* distribution were not robustly different from zero across and within the emotional expression categories (see Table S9 in the Supplemental Materials for all slope comparisons). This suggests that *Confidence in emotion recognition* might not be affected by autistic trait levels.

The same was the case for *Interceptive sensibility*. For neutral expressions, we did, however, find robust evidence for a positive estimated slope at average *Trait interoceptive accuracy*, slope = 0.03, 95% HPD [0.00, 0.06], and “high” (mean + 1SD) *Trait interoceptive accuracy*, slope = 0.03, 95% HPD [0.00, 0.06], with both Highest Posterior Density (HPD) intervals excluding zero. Thus, specifically when evaluating neutral expressions, individuals with a higher *Trait interoceptive accuracy* were more confident in their decisions (see Figure 4E).

Discussion

As expected, we found evidence for a reduced emotion recognition performance with higher autistic trait levels. Yet, only angry facial expressions were significantly less recognized in our experiment. Furthermore, against our hypotheses, higher autistic traits were not associated with a lower trait interoceptive accuracy, both when and when not controlling for alexithymia. Trait interoceptive accuracy was also neither directly linked to emotion recognition accuracy nor had an indirect effect in the association between autistic traits and recognition of angry expressions. The exploratory analyses, however, showed that individuals with higher trait interoceptive accuracy were more confident in judging neutral expressions, and that both neutral and sad expressions were perceived as more emotionally intense with higher interoceptive sensibility. Autistic traits, in contrast, were neither associated with alterations in confidence judgment nor in perceived intensity of emotional expressions.

Taken together, trait interoceptive accuracy did not play a role in altered emotion recognition with higher autistic trait levels, and was not even directly related to emotion recognition outcomes. Yet, interoception might still have an influence on emotion processing as interoceptive sensibility, that is the subjective awareness of bodily signals, affected the perceived intensity of some facial expressions, and trait interoceptive accuracy was linked to confidence ratings associated with recognizing neutral expressions. Being performed in an online setting, we had no insights in an individual’s objective interoceptive accuracy as well as their physiological responses to the facial emotional expressions in Experiment 1. These limitations were addressed in Experiment 2 (see Figure 1).

Experiment 2

Method

Participants

For the replication and extension of Experiment 1, we aimed to have a comparable sample to Experiment 1 (see above for rationale). Thus, 105 participants were recruited, either via the online recruitment platform SONA of Leiden university (97 participants) or via in-person advertisement at Leiden university between 04/01/2022 and 02/05/2023. Of this sample, two participants had to be excluded due to diagnosed neurodevelopmental conditions, two participants because of software failure and one participant because of missing data. Our final sample consisted of 100 participants (88 female, 12 male), aged between 18 and 26 years ($M = 20.02$, $SD = 2.11$). The majority of our participants were Dutch (57 individuals), then German (11 individuals) and Polish (7 individuals), and all participants completed the experiment in English. There was no monetary reimbursement for participation but Leiden university students could receive 3 course credits. Informed consent was provided prior to participation and the study protocol was approved by the local ethics committee of the Faculty of Social and Behavioural Sciences at Leiden University (2022-03-11-M.E.Kret-V2-3838). The experiment was conducted according to the Declaration of Helsinki.

Stimuli

We used the same stimuli as in Experiment 1.

Procedure

The procedure of the lab study was closely matched to the procedure of the online study (see Figure 1). Participants were brought to a quiet experiment room where the facial electromyography (fEMG) recordings were prepared (see Measurements section). All tasks and questionnaires were presented on a Philips screen with a resolution of 1920 x 1080 pixels (23.6") which was at approximately 50cm distance of the participant. The background colour of all tasks was uniform grey. Participants always completed an emotion recognition task first (using EPrime 3.0), which was followed by a heartbeat discrimination task (Whitehead et al., 1977; using PsychoPy 2021.2.3) and all questionnaires in an online survey (using Qualtrics,

see Measurements section). This fixed task order was chosen to avoid biases in the lab study which might not have been present in the online study or might influence task responses. More specifically, the heartbeat discrimination task was always the second task to avoid priming participants to listen to their body signals while performing the emotion recognition task. We also did not want participants to be biased in their responses in the heartbeat discrimination task by activating beliefs about general interoceptive abilities via the questionnaires. Therefore, the questionnaires were always completed last. Experimenters were always present in the room during the instruction and practice phases of each task, and left the room for the main task as well as for the questionnaires.

The trial structure of the emotion recognition task in the lab study was the same as in the online study: After a central fixation cross lasting 1s, a randomly selected facial expressions video (see Stimuli section) was presented in the centre of the screen for 2s (720 x 480 pixels, average visual angle: 22.12° horizontal and 14.85° vertical). A 100ms blank screen was then followed by a screen with questions on the emotion label, on confidence in the emotion label decision and on the perceived emotional intensity of the expression (see Procedure section of Study 1). Upon completion of the emotion recognition task, the fEMG electrodes were removed from the participants' faces, and three electrodes for the electrocardiogram (ECG) were applied to the participants' upper bodies (see Supplemental Materials). To perform the heartbeat discrimination task (Whitehead et al., 1977), participants were given the instruction to judge whether a set of five tones is played "in sync" or "out of sync" with their own heartbeats via key press. Auditory feedback on R-peaks with a delay of 200ms is typically perceived as synchronously by participants with a high cardiac interoceptive accuracy, and a delay of 500ms is perceived as delayed (Ring & Brener, 2018). The usage of multiple delays can provide a better and more individualized estimate of a participant's cardiac interoceptive accuracy (method of constant stimuli (Kleckner et al., 2015)). When piloting the task used in the current study, colleagues found that the interoceptive accuracy index resulting from the two interval method was more closely linked to other measures of interoception than the interoceptive accuracy index from the method of constant stimuli, while the two measures were correlated. As it also requires less time to complete, we decided to use the two interval method for the current study. After a 2min baseline heart rate recording, \ participants judged the synchronicity of five black dots appearing simultaneously (or delayed) with five tones as a

practice (four trials). Each trial of the heartbeat discrimination task started with the visual presentation of numbers counting down from 3 for 3s and a short break (depending on the delay condition), after which participants were presented the five tones via headphones and no visual input (blank grey screen). Once all five tones were played, a question screen appeared asking the participants to judge whether the tones were “in sync” or “out of sync” with their heartbeats by key press. On a second screen, they had to indicate the confidence in their decision on a visual analogue scale from “total guess/no heartbeat awareness” to “complete confidence/full perception of heartbeat”. Integer values ranged from 0-100 but were not visible to the participant. All judgments were made in a self-paced manner. Participants completed 60 trials of the heartbeat discrimination task in total (30 per delay condition). Once the task was completed, the ECG electrodes were removed from the participants’ bodies, and they filled in the questionnaires (see Measurements section) in a randomized order after providing demographical information regarding their age, gender, and nationality.

Measurements

Facial electromyography (fEMG). We used facial electromyography (fEMG) as technique to derive mimicry of the presented emotional expressions. Following the guidelines of Fridlund and Cacioppo (1986), we placed a reusable 4 mm Ag/AgCl surface electrode as a ground electrode on the top of the participants’ foreheads, two electrodes of the same type over the Corrugator Supercilii region (referred to as “corrugator” hereafter) above the participants’ left eyebrows, and two electrodes over the Zygomaticus Major region (referred to as “zygomaticus” hereafter), that is on the participants’ left cheeks. Expressions of sadness, fear and anger are typically associated with increased activations of the corrugator whereas happiness expressions are associated with a decreased activation (i.e., a relaxation) compared to neutral expressions (e.g., Folz et al., 2022; Künecke et al., 2014). Additionally, increased activation over the zygomaticus region typically occurs when happiness is expressed. Thus, facial mimicry of the presented expressions should result in similar muscle activations. The fEMG signal was recorded at a sampling rate of 1000 Hz, using a Biopac MP150 system (see Supplemental Materials for details on the data recording and preprocessing). For each trial, separate epochs were defined for the first 500ms of each video with a neutral expression (as baseline) and the 1.5s in which the emotional expression was shown (as response). Based on an automated detection of extreme values as well as manual coding, 41 trials

of the preprocessed corrugator data (1%) and 149 trials of the preprocessed zygomaticus data (3%) were excluded from further processing. Data of each trial was baseline-corrected by subtraction, z-scored by participant and muscle region and averaged within the response window of each trial.

Electrocardiography. We recorded the participants' electrocardiograms to provide (delayed) auditory feedback about heartbeats in the heartbeat discrimination task (see Procedure section). Three disposable 35mm AG/AgCl electrodes were attached to the participants' upper bodies. The negative electrode (Vin-) was placed under the right collarbone, the positive electrode (Vin+) on the left bottom rib, and the ground electrode below the right ribs. The data was recorded with a sampling rate of 1000Hz using a BIOPAC MP150 system (see Supplemental Materials for details on data recording and preprocessing). As irregularities in the ECG recordings might have resulted in imprecise heartbeat feedback, we visually inspected the recorded data and excluded trials with irregularities from calculating objective interoception measures.

Cardiac interoceptive accuracy and interoceptive trait prediction error. We calculated cardiac interoceptive accuracy by dividing the number of trials that were correctly responded to in the heartbeat discrimination task by the total number of trials (excluding trials with irregularities). To rule out that differences in baseline heart rate could explain individual differences in cardiac interoceptive accuracy, we calculated a correlation (Spearman's rank) between the two measures, which was not significant ($p = 0.69$). The interoceptive trait prediction error was calculated according to Garfinkel et al. (2016): Both cardiac interoceptive accuracy scores and interoceptive sensibility scores were centered and scaled. Then, the difference between the two values was calculated for each participant as a measure of their individual interoception trait prediction error, with positive scores reflecting an overestimation and negative scores reflecting an underestimation of interoceptive abilities. Information about the distribution of both interoception measure scores can be found in Table 3.

Interoceptive attention. We used the Interoceptive Attention Scale (IATS; Gabriele et al., 2022) to assess self-reported interoceptive attention regarding a variety of body sensations (e.g., heartbeat, hunger, need to urinate,...). Interoceptive attention for each of the 21 IATS-items are rated on a 5-point Likert scale (1 = disagree

strongly, 2 = disagree, 3 = neither agree nor disagree, 4 = agree, and 5 = strongly agree). Descriptive information about all questionnaires scores in Experiment 2, including an overview of the questionnaire reliabilities and distribution parameter, can be found in Table 3. A visualization of the relations between the questionnaire measures, as well as the measures from the heartbeat discrimination task, can be found in Figure S1B in the Supplemental Materials.

Interoceptive sensibility. While the same 26 items as in Experiment 1 were used to calculate interoceptive sensibility (i.e., the items of BPQ-SF body awareness scale, see Table 3), participants additionally completed the 20 other items of the BPQ full version body awareness scale in Experiment 2. In addition, we asked the same control question as for the Interoceptive Attention Scale (Gabriele et al., 2022) to unveil whether participants interpreted this scale as a measure of interoceptive accuracy, interoceptive attention or actual frequency and/or intensity of body sensations.

Autistic traits, Social anxiety traits, Alexithymia. To assess autistic traits, social anxiety traits and alexithymia, the same questionnaires were used as in Experiment 1. Descriptive statistics and information about the distribution and reliability are summarized in Table 3.

Table 3. Distribution of questionnaires scores, including their internal consistency (Cronbach's Alpha) and distribution parameters, as well as objective interoceptive accuracy measures in the lab study (Experiment 2)

	N	Mean	SD	Min	Max	α	ICL	hCL	kurtosis	skewness
AQ	100	17.30	5.82	4	38	0.72	0.64	0.80	3.69	0.32
IAS	100	75.64	8.59	53	98	0.74	0.67	0.81	3.21	0.09
BPQ	100	74.16	17.14	39	112	0.91	0.88	0.93	2.24	-0.07
LSAS	100	48.73	20.52	0	120	0.93	0.91	0.95	3.65	0.27
TAS	100	48.25	11.64	24	77	0.84	0.79	0.88	2.60	0.19
IATS	100	57.22	12.54	21	87	0.88	0.84	0.91	2.88	0.09
Cardiac IA	100	0.54	0.08	0.33	0.85				5.31	0.64
ITPE	100	0.00	1.41	-4.94	3.33				3.82	-0.52

Note. AQ = Autism Quotient (Autistic traits), IAS = Interoceptive Accuracy Scale (Trait interoceptive accuracy), BPQ = Body Perception Questionnaire (Interoceptive sensibility), LSAS = Liebowitz Social Anxiety Scale (Social anxiety traits), TAS = Toronto Alexithymia Scale (Alexithymia), IATS = Interoceptive Attention Scale (Trait interoceptive attention), Cardiac IA = Cardiac interoceptive accuracy, ITPE = Interoceptive trait prediction error, ICL/hCL = lower/higher Confidence Level of 95% Confidence Interval associated with α .

Data Analysis

All analyses were preregistered on OSF (<https://osf.io/97a6e>). As explained in the Data Analysis section of Experiment 1, we focused on *Autistic traits* as main predictor in our analyses. Comparable to Experiment 1, significant medium positive correlations between *Autistic traits*, *Social anxiety traits* and *Alexithymia* were observed (LSAS-AQ: $r_s = 0.32, p = .001$; LSAS-TAS: $r_s = 0.25, p = .01$; AQ-TAS: $r_s = 0.35, p < .001$). Prior to model fitting, all continuous variables were standardized (i.e., centered and scaled) to obtain standardized beta coefficients. As a first step, we replicated the mediation analysis as outlined in the Data analysis section of Experiment 1 by fitting three models, using the lmerTest package (Kuznetsova et al., 2017), and quantifying the indirect effect of *Trait interoceptive accuracy* in the association between *Autistic traits* and *Emotion recognition accuracy* for specific *Emotion category* levels using the RMediation package (Tofighi, 2023). We also explored once again whether *Autistic traits*, *Trait interoceptive accuracy* or *Interoceptive sensibility* would be systematically linked to variations in (1) *Perceived emotional intensity* and (2) *Confidence in emotion recognition* in two separate models (see Data analysis section of Experiment 1).

As a second step, we investigated how *Cardiac interoceptive accuracy* would relate to subjective measures of interoception (*Trait interoceptive accuracy* and *Trait interoceptive attention*) and *Autistic traits* by running two zero-order correlation analyses. According to the 2x2 factor model by Murphy and colleagues (2019), we should observe a significant positive relationship between *Trait interoceptive accuracy* and *Cardiac interoceptive accuracy*, whereas there should be no such relationship between *Trait interoceptive attention* and *Cardiac interoceptive accuracy*. Furthermore, a partial correlation between *Autistic traits* and *Cardiac interoceptive accuracy*, while controlling for *Alexithymia*, was performed. Lastly, a potentially stronger *Interoceptive trait prediction error* with higher *Autistic traits* was examined (Garfinkel et al., 2016), using a zero-order correlation. P-values of the four correlations were adjusted with the Holm-method. To test the expected positive relation between *Cardiac interoceptive accuracy* and *Emotion recognition accuracy*, we fitted a binomial GLMM on *Emotion recognition accuracy* (1 = correct, 0 = incorrect) with *Emotion category* (angry, happy, fearful, sad and neutral), *Cardiac interoceptive accuracy* and their interaction as fixed effects, and random intercepts for each participant and each stimulus.

As a third step, we explored whether, with higher autistic trait levels, facial muscle activity would be less predictive of emotion recognition performance for some emotions. To reduce the number of analysis, we decided not to run separate models for each emotion category but to run one model integrating all categories as well as the two muscle regions. More specifically, we fitted a GLMM on *Emotion recognition accuracy* with a three-way interaction between *Emotion category* (angry, happy, fearful, sad and neutral), *Autistic traits* and baseline-corrected, z-scored *Corrugator activity*, a three-way interaction between *Emotion category*, *Autistic traits* and baseline-corrected, z-scored *Zygomaticus activity*, and a two-way interaction between *Emotion category* and *Social anxiety traits*, and *Alexithymia* as control predictors. As in all models, random intercepts for each participant and each stimulus were included.

Results

Replication: Main Analysis Experiment 1

As in Experiment 1, we did not observe a mediation of the effect of *Autistic traits* on *Emotion recognition accuracy* via *Trait interoceptive accuracy* (see Figure 3). The comparison against zero of slopes between *Autistic traits* and *Emotion recognition accuracy* of specific emotions in the first model revealed that only recognition of sad expressions was worse with higher *Autistic traits*, slope = -0.28, 95% CI [-0.54, -0.03], $z = -2.86$, $p = .017$ (**path c**, see Figure 4B, as well as Table S13 in the Supplemental Materials for the full model fit). Therefore, an indirect effect of *Trait interoceptive accuracy* on the association between *Autistic traits* and *Emotion recognition accuracy* of exclusively sad expressions was examined (**path ab**). As confidence limits included zero, $\mu = 0.01$, 95% CL[-0.03, 0.06], we again found no indication that trait interoceptive accuracy would mediate worse emotion recognition with higher autistic trait levels.

Next to a robust negative slope for sad expressions, we also observed a robust positive slope in the relation between autistic trait levels and the recognition of neutral expressions, trend = 0.35, 95% CI [0.08, 0.62], $z = 3.29$, $p = .005$. Unexpectedly, this observation indicates a better recognition of neutral expressions with higher autistic trait levels (see Figure 4B and Table S13 in the Supplemental Materials). In line with Experiment 1, the predictor *Autistic traits* was not significantly linked

to *Trait interoceptive accuracy* in our second model (**path a**, see Table S16 in the Supplemental Materials), and neither *Trait interoceptive accuracy* nor its interaction with *Emotion category* were significant predictors of *Emotion recognition accuracy* (**path b**, see Table S18 in the Supplemental Materials). Again, there was no meaningful change in outcomes if the control predictor *Alexithymia* was excluded from all models with *Autistic traits* as. The model fits of the mediation analysis, including (significant) effects that are unrelated to our predictors of interest, can be found in Tables S13-18 in the Supplemental Materials.

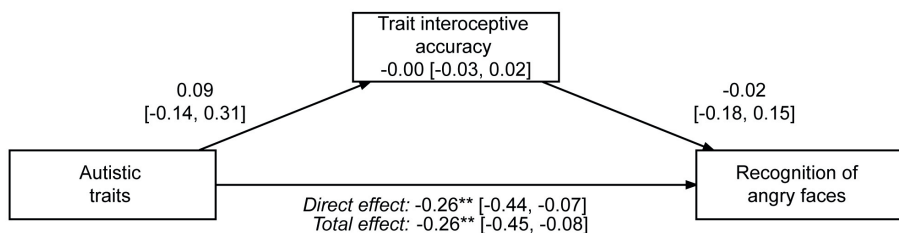
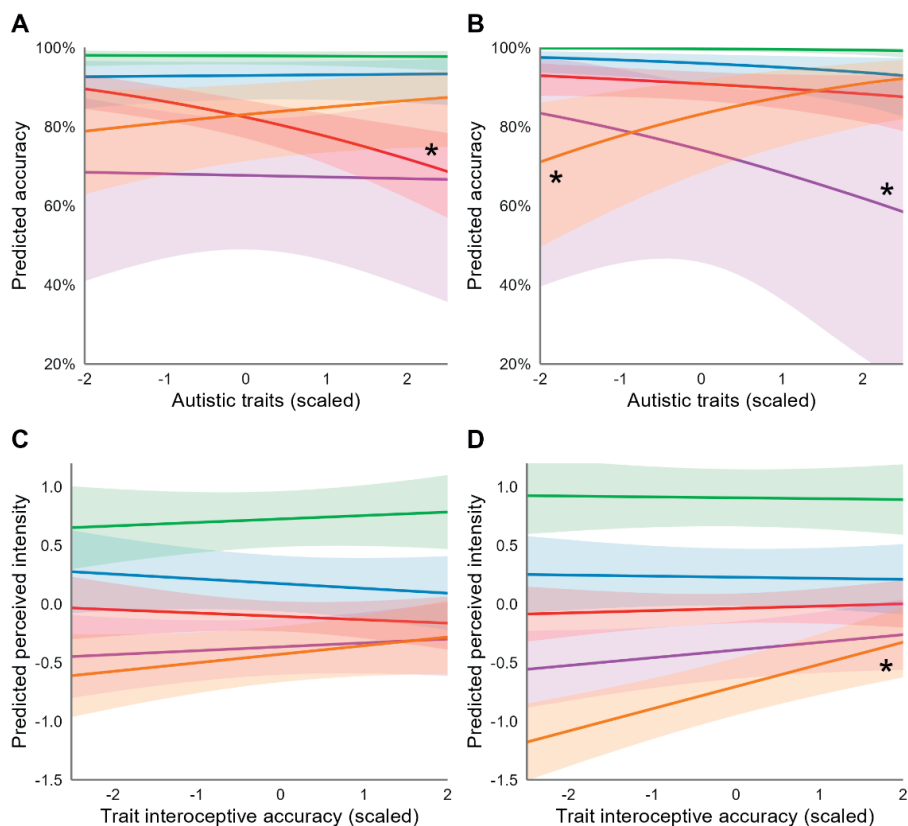


Figure 3. Results of the mediation analysis in Experiment 2 with trait interoceptive accuracy as potential mediator explaining worse recognition of sad faces with higher autistic trait levels. While the coefficient for sad facial expressions in the significant interaction between Autistic traits and Emotion category was not significant, the slope of this effect was robustly negative, which is why we tested for a mediated effect (in line with the main analysis in Experiment 1).



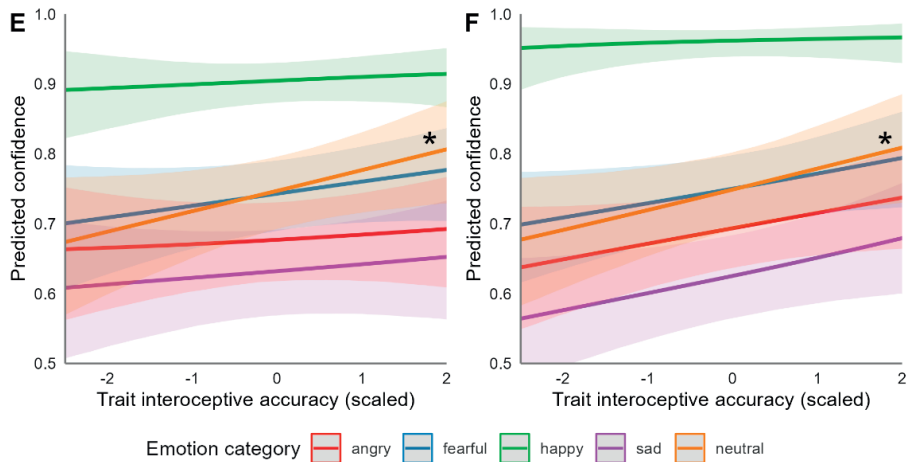


Figure 4. Replicated significant interactions between self-report measures and categories of facial expressions in predicting emotion recognition task outcomes (Experiment 1 on the left and Experiment 2 on the right). Asterisks indicate robust negative/positive slopes for specific emotion categories in the slope comparison against zero. Shaded areas represent 95% confidence boundaries. Predicted accuracy and predicted confidence are on a scale from 0-1, while perceived emotional intensity is centred and scaled, with 0 representing the mean and 1 representing the value at 1 SD.

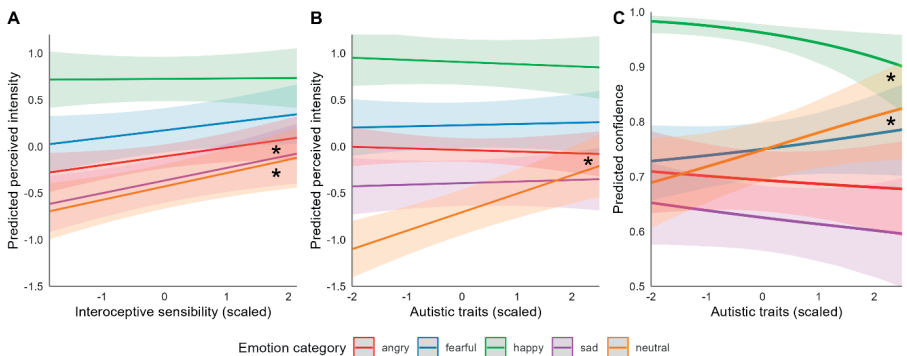


Figure 5. Non-replicated significant interactions between self-report measures and categories of facial expressions in predicting emotion recognition task outcomes in Experiment 1 (A) and Experiment 2 (B+C). Asterisks indicate robust negative/positive slopes for specific emotion categories in the slope comparison against zero. Shaded areas represent 95% confidence boundaries. Predicted accuracy and predicted confidence are on a scale from 0-1, while perceived emotional intensity is centred and scaled, with 0 representing the mean and 1 representing the value at 1 SD.

Replication: Exploratory Analyses Experiment 1

As in Experiment 1, we observed a significant interaction between *Trait interoceptive accuracy* and *Emotion category*, $F(1,4835) = 14.28$, $p < .001$, in the model with *Perceived emotional intensity* as an outcome (see Figure 4D and Table S19 in the

Supplemental Materials for the full model fit). While the linear relations between *Trait interoceptive accuracy* and *Perceived emotional intensity* varied between emotion categories, the slope of the relation was only significantly different from 0 for neutral expressions (see Table 4 for all slope comparisons of significant interactions). In addition, the interaction between *Emotion category* and *Autistic traits* was significant, $F(1,4835) = 17.26, p < .001$ (see Figure 5B). We again observed variation between emotion categories regarding the relation between *Autistic traits* and *Perceived emotional intensity* (see Table S19 in the Supplemental Materials for the coefficients). Neutral expressions were, however, again the only emotion category for which the slope was significantly different from zero (see Table 4). As opposed to Experiment 1, the interaction between *Interoceptive sensibility* and *Emotion category* was not significant in this model (similar results for the model without *Alexithymia* can be found in Table S20 in the Supplemental Materials).

Table 4. Estimated slope of the linear relation between self-report measures and perceived emotional intensity by emotion category, and results of the slope comparisons against zero

Self-report measure		Estimated slope	95% CI	df	t-ratio	p-value
Interoceptive accuracy (IAS)	<i>Emotion category</i>					
	Angry	0.02	-0.10, 0.14	144	0.43	1
	Fearful	-0.01	-0.13, 0.11	144	-0.21	1
	Happy	-0.01	-0.12, 0.11	144	-0.17	1
	Sad	0.07	-0.05, 0.18	144	1.47	0.58
	Neutral	0.19	0.07, 0.31	144	4.23	< .001
Autistic traits (AQ)	<i>Emotion category</i>					
	Angry	-0.02	-0.14, 0.10	140	-0.37	1
	Fearful	0.01	-0.11, 0.13	140	0.28	1
	Happy	-0.02	-0.14, 0.10	140	-0.50	1
	Sad	0.02	-0.10, 0.14	140	0.37	1
	Neutral	0.20	0.00, 0.32	140	4.29	< .001

Note. Confidence intervals are adjusted for multiple comparisons using the Bonferroni-method, and p-values are adjusted using the Holm-method.

In contrast to Experiment 1, estimated slopes at all three examined points of the *Autistic traits* distribution (mean – 1SD, mean, mean + 1SD) were negative and robustly different from zero for happy facial expressions in the Bayesian LMM with *Confidence in emotion recognition* as outcome (see Table 5). This indicates that, across wide parts of the distribution, there was a significant trend for lower

confidence in the recognition of happy expressions with higher *Autistic traits* (see Figure 5C). Moreover, for neutral expressions, we found robust evidence for a positive estimated slope at average *Autistic traits* (see Table 5). Thus, especially at average and “high” autistic trait levels, confidence in the recognition of neutral expressions seems to increase with higher trait levels, matching the findings regarding a better recognition of neutral expressions with higher trait levels. We also observed a significant positive slope at all three examined points of the *Trait interoceptive accuracy* distribution (mean – 1SD, mean, mean + 1SD) across emotions. When splitting by emotional expression categories, the robust positive slope only remained significant for neutral expressions at “low”(mean – 1SD) and average *Trait interoceptive accuracy* (see Table 5). Thus, the effect of higher confidence in emotion recognition with higher trait interoceptive accuracy seems most pronounced in the evaluation of neutral expressions (see Figure 4F). For all other predictors, slopes were not significantly different from 0 across and within emotion categories.

Table 5. Integrated posterior estimates of the Generalized Linear Mixed Model with a zero-one-inflated family for the slopes of the relation between confidence in emotion recognition and self-report measures at three different values (mean -1SD, mean, mean +1SD)

Self-report measure		Mean – 1SD Estimated slope [95% HPD]	Mean Estimated slope [95% HPD]	Mean + 1SD Estimated slope [95% HPD]
Autistic traits (AQ)	<i>Emotion category</i>	0.00 [-0.02, 0.02]	0.00 [-0.02, 0.02]	0.00 [-0.02, 0.02]
	Angry	-0.01 [-0.03, 0.02]	-0.01 [-0.03, 0.02]	-0.01 [-0.03, 0.02]
	Fearful	0.01 [-0.01, 0.03]	0.01 [-0.01, 0.04]	0.01 [-0.01, 0.04]
	Happy	-0.01 [-0.02, -0.00]	-0.02 [-0.03, -0.00]	-0.02 [-0.05, -0.00]
	Sad	-0.01 [-0.04, 0.01]	-0.01 [-0.04, 0.01]	-0.01 [-0.04, 0.02]
	Neutral	0.03 [-0.00, 0.06]	0.03 [0.00, 0.06]	0.03 [0.00, 0.06]
		0.02 [0.00, 0.04]	0.02 [0.00, 0.04]	0.02 [0.00, 0.04]
Interoceptive accuracy (IAS)	<i>Emotion category</i>	0.02 [-0.00, 0.04]	0.02 [-0.00, 0.05]	0.02 [-0.00, 0.05]
	Angry	0.02 [-0.00, 0.04]	0.02 [-0.00, 0.05]	0.02 [-0.00, 0.05]
	Fearful	0.02 [-0.02, 0.04]	0.02 [-0.02, 0.05]	0.02 [-0.02, 0.05]
	Happy	0.00 [-0.01, 0.03]	0.00 [-0.01, 0.02]	0.00 [-0.01, 0.01]
	Sad	0.02 [-0.00, 0.04]	0.03 [-0.00, 0.05]	0.03 [-0.00, 0.06]
	Neutral	0.03 [0.00, 0.06]	0.03 [0.00, 0.06]	0.03 [-0.00, 0.06]
		0.03 [0.00, 0.06]	0.03 [0.00, 0.06]	0.03 [0.00, 0.06]

Note. HPD = Highest Posterior Density interval

Cardiac Interoceptive Accuracy and Self-Report Measures of Interoception, Autistic Traits, and Emotion Recognition Accuracy

Using Mahalanobis distance, we identified and removed bivariate outliers in the relation between *Cardiac interoceptive accuracy* and *Trait interoceptive accuracy* ($n = 7$), *Trait interoceptive attention* ($n = 6$) and *Autistic traits* ($n = 6$), as well as in the relation between the *Interoceptive trait prediction error* and *Autistic traits* ($n = 5$). In line with the theoretical separation between interoceptive accuracy and attention (Gabriele et al., 2022), we did not find a significant relation between *Cardiac interoceptive accuracy* and *Trait interoceptive attention* in our study ($p > .05$). Contrasting our expectations, *Cardiac interoceptive accuracy* was neither positively related to *Trait interoceptive accuracy* nor negatively related to *Autistic traits* (both with and without controlling for *Alexithymia*). There was a trend towards a higher interoceptive trait prediction error with higher *Autistic traits* ($r = .20$, $p = .05$), which did not survive the correction for the four comparisons ($p_{\text{adjusted}} = .20$). The associated correlation matrix can be found in Table S21 in the Supplemental Materials, and a visualization of the relations between all investigated variables Figure S1B in the Supplemental Materials. Lastly, *Cardiac interoceptive accuracy* was not a significant predictor in the GLMM on *Emotion recognition accuracy* ($p > .05$ for both the main effect and the interaction; see also Table S22 in the Supplemental Materials). As *Cardiac interoceptive accuracy* was not related to any of our variables of interest, we did not further investigate its role in *Emotion recognition accuracy* (related to facial mimicry).

Facial Mimicry in Emotion Recognition and its Modulation by Autistic Traits

The model examining whether the link between facial muscle responses and recognition accuracy of distinct facial expressions would be modulated by *Autistic traits* did not reveal any effects beyond those already reported for the first model of the results section (see also Table S23 in the Supplemental Materials). Thus, facial muscle responses were not predictive of *Emotion recognition accuracy*, and there was also no effect of *Autistic traits* on this link.

Exploratory Analysis

Neither interoception measures nor facial muscle activations could explain altered emotion recognition associated with autistic traits in this study, or were themselves significant predictors of recognition accuracy. In line with previous work, trait interoceptive accuracy was significantly linked to the perceived emotional

intensity of some expressions. Interoceptive signals might thus rather alter the representation of experienced and/or observed emotional states, than indicate their qualia. Surprisingly, however, both higher interoceptive accuracy as well as higher autistic trait levels were specifically associated with a higher perceived emotional intensity of neutral expressions in Experiment 2. From an embodied perspective, there could be two potential explanations why observed neutral expressions might be perceived as emotional: either physiological feedback which typically indicates (the lack of) emotionality might not be integrated in the representation of an expression or physiological signals unrelated to the observed expression might be misinterpreted. While, based on previous literature (Garfinkel et al., 2016; Quattrocki & Friston, 2014), the first explanation seems more plausible for the results regarding autistic traits, the second might explain higher perceived emotional intensity of neutral expressions with higher trait interoceptive accuracy (Dunn et al., 2010).

To explore this idea further, we fitted one large model in which we examined whether facial muscle activations would be linked more strongly to perceived emotional intensity with (a) higher trait interoceptive accuracy and (b) lower autistic trait levels. More specifically, we extended the model predicting *Perceived emotional intensity* by adding four three-way interactions, all including *Emotion category* as well as either one of the two facial muscle activations (baseline-corrected, z-scored *Corrugator activity* or baseline-corrected, z-scored *Zygomaticus activity*) and either one of two self-report measures (*Autistic traits* or *Trait interoceptive accuracy*). Next to these predictors of interest, the model still included a two-way interaction between *Emotion category* and *Interoceptive Sensibility*, a two-way interaction between *Emotion category* and *Social anxiety traits*, and *Alexithymia* as control predictors. As in all previous models, random intercepts for each stimulus and each participant were added. The results of this extended intensity model indeed suggest that the directionality of the effect of facial muscle activation on intensity ratings might depend on the trait dimension. More specifically, while both *Corrugator activity* and *Zygomaticus activity* seem to be less predictive of *Perceived emotional intensity* across emotions with higher *Autistic traits*, $\beta = -0.03$, 95% CI [-0.06, -0.01] and $\beta = -0.02$, 95% CI [-0.05, -0.00] (see Figure 6A+B), there was a trend of *Zygomaticus activity* being more predictive of *Perceived emotional intensity* with higher *Trait interoceptive accuracy*, $\beta = 0.02$, 95% CI [-0.00, 0.05] (see Figure 6C, and Table S24 in the Supplemental Materials for the full model fit).

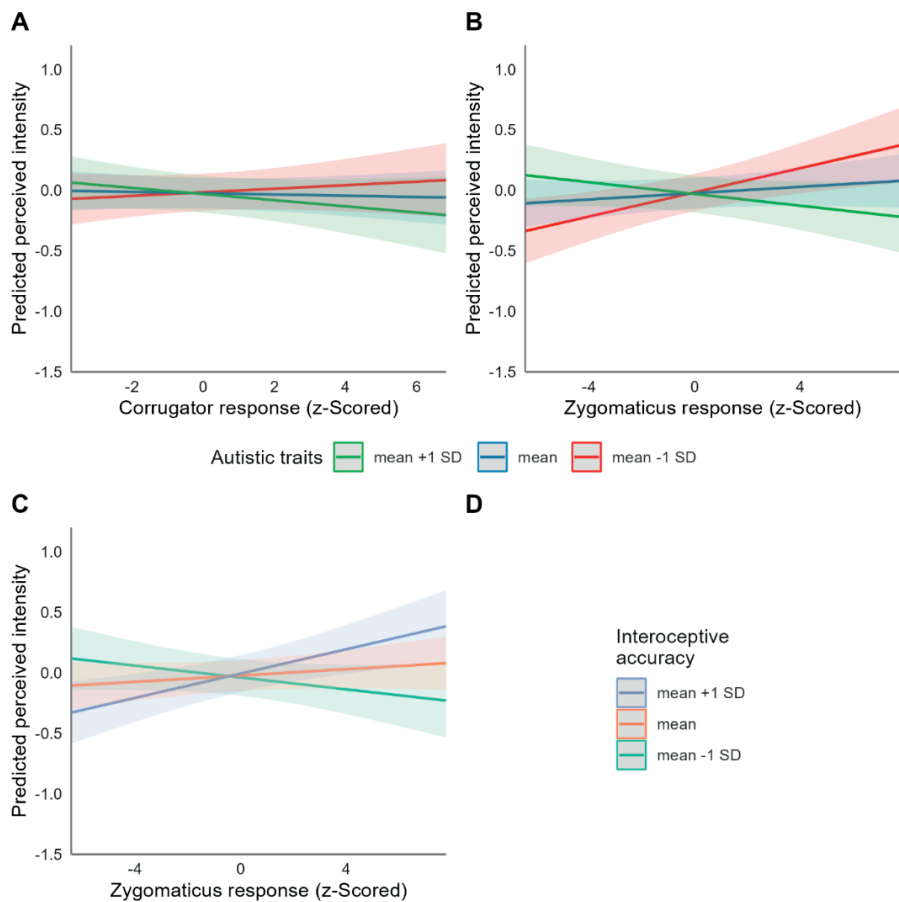


Figure 6. Modulation of the links between facial muscle responses to observed facial expressions and perceived emotional intensity by self-report measures. Continuous self-report measures are split in groups for visualization purposes. Perceived emotional intensity is centered and scaled, with 0 representing the mean and 1 representing the value at 1 SD. Shaded areas represent 95% confidence boundaries.

Sensitivity Analyses

The distributions of both baseline-corrected, averaged and z-scored facial muscle activity signals were highly leptokurtic (kurtosis of 7.65 and 17.83 for corrugator and zygomaticus respectively). To gauge the potential influence of extreme outliers on our findings, we decided to run sensitivity analyses for the models including *Corrugator activity* and *Zygomaticus activity* as predictors. More specifically, we transformed the distribution of the two variables to be as close to Gaussian as possible, using the Gaussianize function of the LambertW R package (Goerg, 2022), and re-ran the last two analyses of this results section. The corresponding model

fits were in line with our original findings: Facial muscle activity remained not predictive of emotion recognition accuracy, and autistic traits did not modulate the relation between the two. Mirroring the results of the exploratory analysis, both gaussianized *Corrugator* activity and gaussianized *Zygomaticus* activity were less predictive of *Perceived emotional intensity* with higher *Autistic traits*, while, with higher *Trait interoceptive accuracy*, gaussianized *Zygomaticus* activity was more predictive of *Perceived emotional intensity* (see Tables S25 and S26 in the Supplemental Materials). Lastly, we aimed to investigate potential systematic effects of response biases in evaluating emotion recognition accuracy in both experiments. Therefore, we calculated each participant's *Unbiased Hit Rates* (i.e., Hu scores; Wagner, 1993) for each emotion category and predicted those by *Autistic traits*, mirroring the first model fits of the mediator models (see Supplemental Materials for more information). Links between *Autistic traits* and *Unbiased Hit Rates* were similar to the results of the main analysis in Experiment 1, whereas there was no evidence for less accurate recognition of sad facial expression with higher Autistic traits in Experiment 2 when looking at *Unbiased Hit Rates*.

Discussion

In the lab study, we could replicate most of the observations that were made in the online study (Experiment 1): First, we did not find evidence for trait interoceptive accuracy being a mediator in the link between autistic trait levels and emotion recognition accuracy, neither being a direct predictor of recognition accuracy. Surprisingly, while accuracy also showed to be reduced with higher autistic trait levels for a specific emotion, this specific emotion was sadness and not anger (as in Experiment 1). Further, recognition of neutral expressions was even increased with higher autistic trait levels in Experiment 2. In line with Experiment 1, higher trait interoceptive accuracy was also linked to a higher perceived intensity of neutral expressions, as well as more confidence in their recognition. In contrast to Experiment 1, we did not find an effect of interoceptive sensibility on perceived emotional intensity. Instead, we found significant effects of autistic traits on the perceived emotional intensity of neutral expressions, as well as on the confidence in recognizing them. Similarly to the results linked to interoceptive accuracy, neutral expressions were rated higher in emotional intensity and confidence in rating them correctly was increased with higher autistic trait levels. Our exploratory

analysis indicated that these seemingly contradicting findings might be the result of integrating actual physiological signals more or less strongly, respectively, in emotional intensity judgments. Confidence in rating happy expressions, in contrast, decreased with higher autistic trait levels. Expanding our models by including objective measures of interoceptive accuracy or physiological changes (i.e., facial muscle responses) in Experiment 2 did not aid to explain emotion recognition accuracy, as well as potential alterations with higher autistic trait levels. Our measure of interoceptive accuracy was also not related to subjective interoceptive accuracy, or any other interoception measure.

General Discussion

Taken the results of our two experiments together, we did not find evidence for either self-reported (trait) or objective (cardiac) interoceptive accuracy explaining modulations in emotion recognition accuracy linked to autistic trait levels. While we did observe lower recognition performance for distinct emotional expressions with higher autistic trait levels, we did not find systematic differences in accuracy linked to individuals' interoceptive abilities. Trait interoceptive accuracy (and interoceptive sensibility in Experiment 1) was rather linked to confidence in emotion recognition as well as the perceived emotional intensity of observed expressions. In contrast to a previous study, facial muscle responses were not predictive of accurately recognizing specific emotional expressions, and the relation between facial muscle responses and emotion recognition accuracy was not altered by an individual's autistic trait levels in the current lab study. Our exploratory analyses, however, indicated that facial muscle activations might be more or less strongly linked to the perceived emotional intensity of an observed expressions, depending on an individuals' trait interoceptive accuracy and autistic trait levels respectively. Thus, as physiological responses and their sensation seem to play a role in altered facial emotion processing in non-autistic individuals with higher autistic trait levels, an examination of their relevance in altered facial emotion processing in autism might yield promising insights.

In line with our expectations, we found some evidence for reduced emotion recognition performance with higher autistic trait levels in our non-autistic sample. Yet, the effect was specific to certain emotion categories and differed between

the two experiments (Experiment 1: anger, Experiment 2: sadness). Inconsistent results with regard to the recognition of distinct facial emotional expressions have also been reported in autistic samples, suggesting specifically worse recognition performance of fear (Uljarevic & Hamilton, 2013), sadness (Wallace et al., 2011), disgust (Enticott et al., 2014; Law Smith et al., 2010), happiness (Eack et al., 2015), or anger (Enticott et al., 2014). While differences in task demands, including stimuli and task complexity, have been suggested as one cause of inconsistencies in autistic samples (Harms et al., 2010), this could not have been the case in the current study, as the same emotion recognition task was performed in both experiments. There were, however, systematic differences in the experimental setting and in sample characteristics, which could have driven the diverging results. Moreover, systematic response biases can distort accuracy scores in categorical judgments (Wagner, 1993), such as on observed emotional facial expressions. In our case, additional analyses with unbiased recognition scores as outcome indicated that these distortions might have played a role in the observed reduced recognition of sad expressions with higher autistic trait levels in Experiment 2, but likely not in the observed reduced recognition of angry expressions (Experiment 1) or the observed increased recognition of neutral expressions (Experiment 2) with higher autistic trait levels (see Supplemental Materials).

This increased recognition of neutral facial expressions with higher autistic trait levels in Experiment 2 unexpected. Perceiving emotionality in neutral facial expressions is a commonly made mistake in the general population, and this bias has been associated with the importance of facial emotion perception in navigating our social world (Albohn Daniel N. and Brandenburg, 2019). By following a rule-based path to facial emotion recognition, which rather relies on the presence/absence of distinct features (Rutherford & McIntosh, 2007), one might be less prone to incorrectly interpret emotionality into neutral faces (i.e., less “false alarms”). Individuals on the autism spectrum have been suggested to follow this rule-based path to emotion recognition, as might non-autistic individuals with high autistic trait levels. The interpretation of facial expressions could further be facilitated by an accurate sensation of physiological signal changes in an emotion recognition context. Based on interoception research in autism (DuBois et al., 2016), we assumed that this path to emotion recognition might be less reliable in non-autistic individuals with higher autistic trait levels. Against our expectations, we did not find that interoceptive accuracy predicted emotion recognition accuracy,

or mediated the negative relation between autistic trait levels and recognition of specific facial expressions. Facial muscle responses were not less predictive of emotion recognition accuracy with higher autistic trait levels as found in a previous study. While our initial hypotheses were not supported in the current study, the findings, nevertheless, offer further insights into the conceptualisation of interoception, its potential role in facial emotion perception as well as putative interoceptive alterations associated with autistic trait levels within this context.

Contradicting previous findings (Murphy et al., 2020), cardiac interoceptive accuracy, measured via the heartbeat discrimination task, was not associated with trait interoceptive accuracy, measured via the interoceptive accuracy scale, in the current study. The use of classic cardiac interoception tasks to assess interoceptive accuracy has recently been heavily criticized. They are likely to be confounded by the reliance on additional cognitive processes (Hickman et al., 2020; Ring & Brener, 2018) and only focus on one interoceptive domain (i.e., cardioception) as well as few dimensions of interoception, neglecting its multidomain and multidimensional nature (Suksasilp & Garfinkel, 2022). The few previous studies comparing the performance of individuals on the autism spectrum to non-autistic controls in heartbeat discrimination tasks did not find a robustly reduced interoceptive accuracy (Z. J. Williams et al., 2023). We did not find a significant negative relationship between cardiac interoceptive accuracy and autistic trait levels in our non-autistic sample either, and, unexpectedly, also no relations between self-reported measures of interoception (i.e., accuracy and sensibility) and autistic traits. Importantly, the expectations on links between interoception measures and autistic trait levels in our non-autistic sample were based on results of studies with individuals on the autism spectrum, who might differ in their interoceptive processing. Alexithymia, in contrast, showed the expected negative relation to trait interoceptive accuracy. While recent perspectives claim that alexithymia might underlie emotion recognition difficulties in autism, our findings on links between autistic traits and all emotion recognition task outcomes in a non-autistic sample did, not change depending on whether alexithymia was included as control predictor or not. Whether alexithymia does play a role in altered facial emotion recognition with higher autistic trait levels in non-autistic individuals, or whether it might only become relevant in autism is to be further investigated.

Higher interoceptive accuracy has been associated with more flexibility in the precision of interoceptive prediction errors (i.e., the potential to increase precision via attention) in the predictive coding perspective, whereas extremely precise, inflexible priors have been proposed as source of interoceptive difficulties in autism (Quattrocki & Friston, 2014). These alterations in priors might already, to some degree, be present in non-autistic individuals with high autistic trait levels, as other phenomena that are associated with autism. Our findings regarding the differential effect of trait interoceptive accuracy and autistic traits on the link between facial muscle activity and perceived emotional intensity could be interpreted as support for these assumptions: In individuals with high interoceptive accuracy, prediction errors associated with interoceptive signals reflecting facial muscle activations during facial expression observation could have been more influential due to a higher precision via attention. As a consequence, they might have reported perceiving stronger emotional experiences (and potentially also showed stronger physiological reactions), which generalized to neutral expressions. Scoring higher on autistic traits, in contrast, might have been associated with less integration of interoceptive information about facial muscle activations on high-level processing, i.e., in attributing emotionality to observed facial expressions. In those lines, higher confidence in the recognition of neutral expressions might be the product of different processes in relation to autistic traits versus trait interoceptive accuracy. With higher autistic traits, recognition of neutral expressions was indeed better, which could relate to a rule-based path to emotion recognition: Not identifying clear indicators of a specific emotion category could have resulted in higher confidence ratings regarding the categorization of neutral expressions, while undefined physiological arousal could have still resulted in rating them as more intense. Lower confidence in recognizing happy expressions, in contrast, might result from integrating mimicry responses less, which are usually most pronounced for happy facial expressions. General higher confidence ratings with higher trait interoceptive accuracy, and specifically for neutral expressions, could be a result of more strongly perceived physiological feedback which reinforced confidence in all labelling decisions (similar to perceived emotional intensity). These are, however, only assumptions which should be tested more systematically in future studies.

Both autistic trait levels as well as interoceptive accuracy modulated the link between instant facial muscle activations and perceived emotional intensity

of expressions. This indicates that the integration of one's own physiological feedback in processing other's emotions seems to differ between individuals, and specifically seems to be reduced with higher autistic trait levels in the general population. Training one's interoceptive abilities might thus be useful to gain emotional clarity not only about one's own but also about others' emotions. The effect of enhancing interoceptive processing with different interventions has already been investigated with varying complexity and on multiple time scales. Priming an interoceptive focus in the general population, for example, has been shown to improve emotion recognition when using a heartbeat counting task (Salamone et al., 2021), but not when using instructions (Bornemann et al., 2012). Addressing interoceptive processing more broadly and on a longer time-scale, (brief) mindfulness interventions have been shown to have a small but significant effect in reducing negative affect in non-clinical and clinically diverse samples (Schumer et al., 2018). For example, online mindfulness interventions have been found to be similarly successful as online cognitive behaviour therapy in reducing anxiety symptoms in some individuals on the autism spectrum (Gaigg et al., 2020). The novel Aligning Dimensions of Interoceptive Experience (ADIE) therapy, which specially focusses on addressing the mismatch between cardiac interoceptive sensibility and accuracy (Quadt et al., 2021), confirms the potential of interoception-based interventions to alleviate anxiety in individuals on the autism spectrum. While the effectiveness of interventions targeting interoception has mainly been investigated in relation to anxiety in autism, interoception also plays a fundamental role in representing the self versus the other (Gao et al., 2019; Palmer & Tsakiris, 2018), including emotional states (Engelen et al., 2023). Given that altered self-other distinction has been linked to social difficulties in autism (Lamm et al., 2016; Lombardo et al., 2010), the usefulness of interoception interventions in strengthening self-other knowledge, including the affective domain, should be examined in future research.

The main objective of our study was to get a better understanding of the role of interoception and physiological responses (i.e., facial mimicry) in emotion processing in relation to autistic trait levels. By closely matching the lab study (Experiment 2) to the preceding online study (Experiment 1), we successfully unveiled robust modulations in the associations between some self-report measures and emotion recognition outcomes. Yet, not all findings could be replicated in the lab study, and some novel observations were made. This could be

explained by differences in the experimental setting and the addition of measures (i.e., questionnaires, physiological recordings and tasks). Furthermore, the sample in Experiment 1 was likely more varied in terms of education and lifestyle than the sample in Experiment 2. Thirty percent of the participants in Experiment 1 were not directly recruited at Leiden university whereas we exclusively recruited at Leiden university for Experiment 2. Participants in both experiments were, nevertheless, majorly female young adults, which limits the generalizability of our results. Age and gender differences have been reported with regard to emotion perception and interoception measures (Grabauskaitė et al., 2017), next to gender differences in autistic traits (Ruzich et al., 2015). Future experiments should therefore not only examine whether our findings could be extended to a clinical population (i.e., individuals with an autism spectrum diagnosis), but also more diverse non-autistic samples. While limitations of the heartbeat discrimination task have already been addressed earlier, computerized emotion recognition tasks can never fully reflect emotion recognition in a naturalistic context. By adding task demands to the perception of facial expressions, spontaneous reactions might be altered and processing of expressions might be biased. An experimental context, which requires labelling of facial expressions and already provides categories for it, might activate a top-down mode and reinforce a specific path to emotion recognition, such as visual matching with mental representations (Keating & Cook, 2023). Bodily responses might thus not act as simulations of specific emotional states but rather reinforce perceived emotionality and confidence in one's decisions in people with higher self-reported interoceptive abilities. Whether interoception could become a relevant factor in emotion recognition in daily life, and whether the path to emotion recognition involving interoception is less commonly used in individuals with higher autistic trait levels (or on the autism spectrum), still needs to be further investigated with more ecologically valid and naturalistic paradigms.

Many different paths can lead to recognizing another person's emotion based on their facial expression. Feedback from our own bodies might be one of them, which might be less strongly pronounced in autism. In the current study, we did not find support for reduced interoceptive accuracy explaining worse emotion recognition performance with higher autistic trait levels. Nevertheless, higher trait interoceptive accuracy (and sensibility) resulted in more confidence in labelling expressions as well as in higher perceived emotional intensity ratings of expressions with little or no visual indication of emotionality. Based on the

outcomes of an exploratory analysis, we suggest that this might be associated with a stronger integration of actual physiological signals indicating an emotional experience. Higher autistic trait levels, in contrast, might be associated with less integration of physiological signals in determining the emotional intensity of an observed expression. These suggestions should be carefully tested in future studies. With more and more research highlighting the role of interoception in experiencing emotions and connecting to others, fostering adaptive interoceptive processing might become an important avenue in facilitating everyday life social encounters in individuals on the autism spectrum.