Adaptive face coding contributes to individual differences in facial expression recognition independently of affective factors

Romina Palermo\textsuperscript{1,2*}, Linda Jeffery\textsuperscript{1,2*}, Jessica Lewandowsky\textsuperscript{1,2}, Chiara Fiorentini\textsuperscript{1,3}, Jessica L. Irons\textsuperscript{3}, Amy Dawel\textsuperscript{1,3}, Nichola Burton\textsuperscript{1,2}, Elinor McKone\textsuperscript{1,3} & Gillian Rhodes\textsuperscript{1,2}

\textsuperscript{1} ARC Centre of Excellence in Cognition and its Disorders
\textsuperscript{2} School of Psychological Science, The University of Western Australia, Crawley, WA, 6009, Australia.
\textsuperscript{3} Research School of Psychology, The Australian National University, Canberra, ACT, 0200, Australia.

* These authors contributed equally.

Corresponding Author:
Associate Professor Romina Palermo
ARC Centre of Excellence in Cognition and its Disorders, and School of Psychology, University of Western Australia, 35 Stirling Highway, Crawley, WA, 6009, Australia
tel: +61 (0) 8 6488 3256
e-mail: romina.palermo@uwa.edu.au

Notes:
Jessica Irons is now at the Department of Psychology, The Ohio State University and Chiara Fiorentini is now at the Swiss Center for Affective Sciences, University of Geneva, Switzerland.

Running head:
Individual differences in facial expression recognition
Abstract

There are large, reliable individual differences in the recognition of facial expressions of emotion across the general population. The sources of this variation are not yet known. We investigated the contribution of a key face perception mechanism, adaptive coding, which calibrates perception to optimise discrimination within the current perceptual “diet”. We expected that a facial expression system that readily recalibrates might boost sensitivity to variation among facial expressions, thereby enhancing recognition ability. We measured adaptive coding strength with an established facial expression aftereffect task and facial expression ability with three tasks optimised for the assessment of individual differences. As expected, expression recognition ability was positively associated with the strength of facial expression aftereffects. We also asked whether individual variation in affective factors might contribute to expression recognition ability, given that clinical levels of such traits have previously been linked to ability. Expression recognition ability was negatively associated with self-reported anxiety but not with depression, mood, or degree of autism-like or empathetic traits. Finally, we showed that the perceptual factor of adaptive coding contributes to variation in expression recognition ability independently of affective factors.
Statement of the public significance of the work

Recognising facial expressions of emotion is an important skill. Our study examined why people vary in their ability to recognise facial expressions. It has been suggested that adaptive coding, a perceptual factor that recalibrates the perceptual system based on recent visual experience, helps differentiate between expressions, and may be associated with better facial expression recognition ability. We asked whether people who showed stronger adaptive coding of facial expressions would show better recognition. We found that they did. We also found that people who self-reported more anxiety were poorer at recognizing facial expressions. Interestingly, adaptive coding and anxiety affected individual differences in expression recognition ability independently. These results suggest that there are multiple, independent reasons why people vary in their ability to recognise facial expressions. The existence of multiple factors is a challenge for researchers designing interventions to improve facial expression recognition.
Adaptive face coding contributes to individual differences in facial expression recognition independently of affective factors

Facial expressions of emotion can communicate moods and intentions, and the ability to accurately perceive, recognise and understand the facial expressions of others is crucial for successful social interactions. Although the human face can display a wide variety of facial expressions, emotion research has tended to focus on the recognition of a set of six so-called ‘basic’ facial expressions (happiness, anger, disgust, sadness, fear and surprise) that are each characterised by a specific set of facial movement patterns (Ekman, 1992). Deficits in recognising these basic facial expressions are often apparent in people with developmental or neurological conditions, such as autism (Uljarevic & Hamilton, 2013), Parkinson’s disease (Marneweck, Palermo, & Hammond, 2014), psychopathy (Dawel, O’Kearney, McKone, & Palermo, 2012), schizophrenia (Kohler, Walker, Martin, Healey, & Moberg, 2010), amygdala damage (Palermo, Schmalzl, Mohamed, Bleasel, & Miller, 2010) or lesions to the orbitofrontal cortex (Willis, Palermo, McGrillen, & Miller, 2014). For people with these clinical conditions, difficulties recognising facial expressions are commonly evident when their performance is compared to the average performance of groups of ‘typical’ individuals in the general population, who report neither developmental or neurological disorders nor significant brain injuries.

However, it is important to note that there is considerable individual variation in the performance of these typical individuals (Palermo, O’Connor, Davis, Irons, & McKone, 2013; Wilhelm, Hildebrandt, Manske, Schacht, & Sommer, 2014). Given the critical role of expression recognition ability in social interactions, it is likely that this variation among typical individuals has an impact on social functioning. Much research has been conducted on individual differences in emotion perception (e.g., Hamann & Canli, 2004; Japee, Crocker, Carver, Pessoa, & Ungerleider, 2009). Nevertheless, the sources of this variation in
expression recognition ability have not yet been established. In the present study we investigated perceptual and affective factors that might contribute to individual variation in facial expression recognition ability.

**A perceptual factor that might be related to facial expression recognition ability**

A key perceptual mechanism of face perception that could underlie this individual variation in typical expression recognition ability is adaptive coding, in which the appearance of faces is affected by the appearance of faces seen previously. Adaptive coding is demonstrated by aftereffects, in which perception is biased away from recently viewed stimuli. That is, exposure (adaptation) to a facial expression (e.g., happy) biases perception away from that expression, producing an expression aftereffect (e.g., a subsequently viewed face appears less happy) (Burton, Jeffery, Calder, & Rhodes, 2015; Burton, Jeffery, Skinner, Benton, & Rhodes, 2013; Cook, Matei, & Johnston, 2011; Hsu & Young, 2004; Skinner & Benton, 2010; Skinner & Benton, 2012; Vida & Mondloch, 2009; Webster, Kaping, Mizokami, & Duhamel, 2004). Expression aftereffects likely tap high-level visual processes involved in expression perception because adaptation occurs in high-level visual areas (e.g., Xu, Liu, Dayan, & Qian, 2012), expression aftereffects transfer across changes in position, viewpoint and size (Benton et al, 2007; Burton et al, 2013) and aftereffects reflect adaptation of both facial features and configurations (Butler, Oruc, Fox, & Barton, 2008).

Adaptive coding is a ubiquitous feature of perceptual coding and has been argued to play a functional role in perception by calibrating perception to optimise sensitivity to prevailing variation and maximise sensitivity to novel input (see Thompson & Burr, 2009, for an outline). It therefore seems plausible that individual variation in the operation of these calibration processes might contribute to discrimination ability. Evidence supporting this proposal has been found for face identity recognition where individual variation in the strength of adaptive coding is positively associated with face recognition ability in typical
individuals (Dennett, McKone, Edwards, & Susilo, 2012; Rhodes, in press; Rhodes, Jeffery, Taylor, Hayward, & Ewing, 2014). Further evidence consistent with a functional role for adaptive coding comes from studies showing that adaptive coding of face identity is reduced in clinical groups with known face recognition deficits such as autism (Pellicano, Jeffery, Burr, & Rhodes, 2007; Rhodes et al, 2014), congenital prosopagnosia (Palermo, Rivolta, Wilson, & Jeffery, 2011) and cataract removal patients (Rhodes, Nishimura, de Heering, Jeffery, & Maurer, 2017).

Studies of individual differences and special populations assume that individual variation in adaptability has long-term effects on coding in the visual system, and that a sample of the system’s adaptability can be measured by strength of aftereffects. However, other types of studies have focused on potential benefits of adaptation over short time scales. For instance, participants may be adapted to a category of faces (e.g., Caucasian faces), with the prediction that subsequent discrimination performance would be enhanced for faces from the adapted category. Although some early studies failed to find improvements (e.g., Dakin & Omigie, 2009; Ng, Boynton, & Fine, 2008; Rhodes, Maloney, Turner, & Ewing, 2007), later studies have shown that adaptation can enhance discrimination around the adapted face (Chan, Yang, Wang, & Fang, 2010; Oruç & Barton, 2011; Rhodes, Watson, Jeffery, & Clifford, 2010; Yang, Shen, Chen, & Fang, 2011).

The relationship between adaptability and face recognition ability has been interpreted within a norm-based coding framework, in which faces are visually represented in a multidimensional space by how they deviate from an average, or “norm”, face that has average values on each dimension in the space (Rhodes et al., 2005; Robbins, McKone, & Edwards, 2007). Crucially, this norm is not fixed but rather is constantly updated by visual experience (adaptation) to match the current “diet” of faces and thereby optimize the ability to discriminate among individual faces (Clifford & Rhodes, 2005; Rhodes & Leopold, 2011; Webster & MacLeod, 2011). This calibration is reflected in face aftereffects where the norm
is temporarily shifted toward to the adapting face and perception is therefore biased toward faces opposite the adaptor because these now deviate more from the norm (Leopold, O’Toole, Vetter, & Blanz, 2001; Rhodes & Jeffery, 2006). Face aftereffects typically decay rapidly. However, the rate of decay is linked to the duration of adaptation (Leopold et al, 2005; Rhodes et al, 2007, Burton et al, 2016) and some effects can be surprisingly enduring (Carbon et al, 2005, 2006, 2007, 2013; Kloth & Rhodes, 2016). It is therefore plausible that consistent visual experience over longer durations (days to years) could result in long-lasting changes to the norm (for Discussion see Carbon et al 2011, Webster 2015). Having a system that adapts readily may be beneficial because it allows for greater tuning of the norm by experience.

Facial expression also appears to be coded relative to a norm in a multidimensional face space, with this norm likewise being updated by experience (Burton et al., 2013; Cook Matei, & Johnson, 2011; Skinner & Benton, 2010). Support for norm-based coding of expression comes from studies in which anti-expressions were created by taking a basic expression and morphing it through a norm (average of all the basic expressions) into the other-side of face-space to create a stimulus that is the opposite of the basic expression (see Figure 1). Adapting to the anti-expression creates an aftereffect, in which perception of subsequently viewed faces is biased toward the opposite expression (e.g., resembling happy after adapting to anti-happy). Though this finding alone is insufficient to rule out alternative, non norm-based models, several studies have produced findings those models have difficulty accounting for, yet are readily accommodated by a norm-based account (Rhodes et al, 2016; Burton et al, 2013; 2015; Skinner & Benton, 2010, 2012). Overall, these results provide good support for norm-based coding of expression. It is plausible, therefore, that efficient updating of face norms may improve sensitivity to expressions, resulting in greater ability to distinguish between facial expressions, to detect subtle expressions and possibly to better detect dynamic changes in expression.
In the present study, we investigated whether adaptive coding of expression is associated with the ability to recognise facial expressions of emotion. If adaptation plays a functional role in expression coding we would expect to observe a positive relationship between the strength of adaptive coding of expression and individual differences in the ability to recognise expression. In a recent study, we found preliminary evidence for such a relationship, using a brief adaptive expression coding task and a single test of expression labelling (Rhodes et al., 2015). Here, we use a more robust design, which includes more than one measure of expression recognition ability, to better measure individual differences in this ability. We measured adaptive coding ability using a longer (double the number of trials) and more reliable expression aftereffect task, as described in Burton et al. (2013), where participants adapt to anti-expression adaptors and then judge the expression of an expression-neutral test face as one of four basic expressions (happy, angry, sad, scared). The degree to which perception of this test face is biased by the adapting anti-expression provides a measure of the strength of adaptive coding.

Performance on aftereffect tasks can be affected by attention, with better attention leading to larger face identity aftereffects, for example (Rhodes, Jeffery, et al., 2011). Variations in the strength of aftereffects due to variations in attention could either add noise to our measure of adaptability, or inflate correlations, if there are consistent individual variations in attention that apply to all tasks. Rhodes et al. (2015) encouraged attention to the adapting faces by presenting a small asterisk on the screen for 150 ms in one of the interstimulus intervals on half the trials and asking participants to report when it was present, at the end of the trial. Here, we used a different attention manipulation, devised by Burton et al. (2013), to encourage good attention each of the adapting faces. Specifically, the eyes or lips of some adaptor faces in each sequence were brightened and participants reported whether the change was to the eyes or mouth via a key press (see Method). This task also provided a measure of attention that we could use to check whether any
association between the strength of expression aftereffects and expression ability could be accounted for by variation in attention to the adaptors.

Figure 1. Stimuli from Skinner and Benton (2010) and Burton et al. (2013) used in this study. The average expression is a blend of six basic expressions (happy, sad, angry, fearful, surprised, disgusted) and neutral. Anti-expressions were formed by morphing an original basic expression (100% Expression) along a vector through the average face into the opposite side of face-space. The 100% anti-expressions were the direct opposite of each expression. The 50% anti-expressions lie halfway between the average and 100%.


Affective factors that might be related to facial expression recognition ability

A secondary aim of the present study was to assess the relative contribution of our perceptual factor to expression recognition ability, that is, in comparison to that of other potential contributors. To do so, we also measured individual differences in affective factors that have been previously associated with expression ability in clinical and also sometimes non-clinical populations. This allowed us to compare the contribution of these affective factors to any contribution of adaptive coding. The affective factors we measured were
depression, anxiety, mood, autism-like and empathetic traits. Adults who exhibit clinical levels of depression and anxiety typically show reduced facial expression recognition performance (see meta-analysis and reviews by Bistricky, Ingram, & Atchley, 2011; Demenescu, Kortekaas, den Boer, & Aleman, 2010), but results from non-clinical samples are more equivocal. For instance, high- and low-trait anxiety groups did not differ in their recognition of basic facial expressions, except for the recognition of fearful faces, in which the high-trait anxiety group actually performed better (Surcinelli, Codispoti, Montebarocci, Rossi, & Baldaro, 2006). Mood also affects the recognition of basic emotions, with the inducement of a sad mood in typical adults reducing performance (Chepenik, Cornew, & Farah, 2007). Autism-like traits are also linked with expression recognition: Typical individuals reporting more autism-like traits were less accurate at recognising expressions than those reporting few autism-like traits (for anger, disgust and sadness but not fear, happiness or surprise, Poljac, Poljac, & Wagemans, 2013) and autism-like traits have been negatively associated with expression recognition ability across the general population (Lewis, Lefevre, & Young, 2016; but see Halliday, MacDonald, Sherf, & Tanaka, 2014). In contrast, because empathy involves the capacity to experience and comprehend the emotions of others (Jolliffe & Farrington, 2006), a positive association with expression recognition ability is expected. Positive associations have been reported between self-reported empathy (or at least some sub-types of empathy) and facial expression recognition ability (Besel & Yuille, 2010; Lawrence, Shaw, Baker, Baron-Cohen, & David, 2004; Lewis et al., 2016).

We measured individual differences in these affective factors using well-established self-report measures. Anxiety and depression levels were estimated with the 21-item version of the Depression Anxiety and Stress Scales (DASS) (Lovibond & Lovibond, 1995). Mood was indexed with the Positive and Negative Affect Schedule (PANAS) (Watson, Clark & Tellegen, 1988), where high positive affect is indicative of enthusiasm and alertness whereas
high negative affect reflects lethargy and sadness. The Autism Quotient (AQ) (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001) is appropriate to measure autism-like traits across the general population (Ruzich et al., 2015) and was used here. To evaluate empathy we administered the Empathy Quotient (EQ) (Baron-Cohen & Wheelwright, 2004; Lawrence, et al., 2004), which measures a single dimension of empathy (Allison, Baron-Cohen, Wheelwright, Stone, & Muncer, 2011) and the Basic Empathy Scale (BES) (Jolliffe & Farrington, 2006), which measures both cognitive and affective empathy.

**Measuring expression recognition ability**

We measured emotion recognition ability comprehensively with a battery of three newly developed, valid and reliable tests that are sensitive to individual differences in ability and measure partially overlapping, but not identical, aspects of emotion recognition: an emotion-matching task (Palermo et al., 2013), an emotion-labelling task (Palermo et al., 2013) and the Multimodal Emotion Recognition Test (MERT) (Bänziger, Grandjean, & Scherer, 2009). In the emotion-matching task, participants judge which one of three simultaneously presented faces displays a different expression, and so this task taps expression discrimination without requiring explicit identification, which can be affected by the provision of categorical labels (Palermo et al., 2013). In contrast, the emotion-labelling task taps explicit identification by requiring participants to select which one of six ‘basic’ emotion labels (anger, disgust, happiness, fear, sadness, surprise) best reflects the expression displayed by each face (Palermo et al., 2013). We also included the MERT (Banziger et al., 2009), which assesses the ability to recognise both static and dynamic faces, with and without emotional vocalisations (and emotional vocalisations alone). The MERT also includes a wider range of expressions (hot anger, cold anger, panic fear, anxiety, despair, sadness, elation, happiness, contempt, disgust) than the matching and labelling tasks. Expression recognition ability was indexed by percentage correct on each task. In
addition, we combined scores across the three tasks using principal components analysis (PCA) to derive a measure of general ‘emotion recognition’ for each individual, independent of task.

In sum, we measured expression recognition ability with a suite of three tasks and examined whether performance on these tasks is linked with a face-sensitive perceptual mechanism (adaptive norm-based coding of expression) and/or individual differences in affective traits (levels of depression, anxiety, mood, autism-like and empathetic traits). Furthermore, we compared the relative size of contributions and investigated whether contributing factors are independent or interrelated.

Method

Participants

The Human Research Ethics Committees of the University of Western Australia (UWA) and the Australian National University (ANU) granted ethical approval. Data from 88 adults of European descent (38 men) (M = 19.92, SD = 3.99, 17-44 years) were analysed (46 were tested at UWA). An additional 4 participants were excluded from all reported analyses: n = 1 due to reported very poor visual acuity in one eye; n = 3 due to scoring below the chance level of 50% on the attention control task included in the expression aftereffect task, suggesting that they did not understand, or were unwilling to follow, instructions. Participants were reimbursed with either course credit or $20.

Tasks and measures

In an effort to measure genuine individual differences in emotion recognition ability, participants completed three emotion-recognition tasks. They also completed an expression aftereffect task, to measure individual differences in adaptive coding of expression, a
perceptual factor that we expect to be associated with emotion recognition ability. Affective factors that might be linked with emotion recognition ability were assessed with a suite of self-report measures, all of which are described below.

**Emotion-recognition tasks**

1. *Emotion-matching task* (Palermo et al., 2013)

   We used the 100-item task described in Palermo et al. (2013). This task uses full-colour images of Caucasian males and females displaying six basic facial emotions (happy, sad, angry, surprised, disgusted and fearful) from the Karolinska Directed Emotional Faces database (KDEF, Lundqvist, Flykt, & Öhman, 1998) in full-face and three-quarter poses, enclosed in an elliptical grey oval that excluded the hair. On each trial, three different individuals of the same sex and with the same pose were shown so that two distractors displayed the same emotion (e.g., anger) and the target displayed a different ‘odd-one-out’ emotion (e.g., disgust). The facial emotions displayed by the target and distractors were paired to be maximally confusable (as per the Emotion Hexagon; Young, Perrett, Calder, Sprengelmeyer, & Ekman, 2002). Participants used numbered keys to select which face (1, 2, or 3) displayed the ‘odd-one-out’ emotion, either while the faces were presented (4,500 ms) or during an additional 7 sec time window after the faces had been erased. The 100 trials were presented in the same order for each participant, and were preceded by eight practice trials.

2. *Emotion-labelling task* (adapted from Palermo et al., 2013)

   We used the 100-item task from Palermo et al. (2013) but made two modifications designed to improve the test’s ability to measure individual differences in expression labelling. First, we reduced the presentation time for each face from 1,000 ms to 400 ms, to increase the overall difficulty of the task and minimise ceiling effects, thereby providing a
better estimate of individual differences in ability. Second, due to concerns that unequal numbers of faces in each emotion category might affect labelling we equated the number of facial expressions in each emotion category so that there were 24 each of angry, disgusted, fearful, happy, sad, and surprised. To do this, we added in 44 new faces from the KDEF. Participants viewed each individual face and used a computer mouse to select the appropriate emotion label from a set of six presented underneath the face. Responses were possible while the face was presented or during a subsequent 7 sec window in which the labels remained on the screen. The 144 trials\(^1\) were presented in the same order for each participant, and were preceded by four practice trials.


This task is described in detail in Bänziger et al. (2009). It was run from the Swiss Center for Affective Sciences server, with instructions translated into English. The MERT is based on short video clips of actors displaying facial emotions and reciting nonsense phrases in different emotional tones. The stimuli were presented in four different modes: (i) audio and video combined, (ii) video only, (iii) audio only, and (iv) still images from the videos. Headphones were provided and participants were able to adjust the sound to the level they desired. Participants selected which one of ten labels presented next to the face best described the stimulus (anxiety, panic/fear, happiness, elation, cold anger, hot anger, sadness, despair, disgust, and contempt). One hundred and twenty trials were presented in one of two different random orders.

\(^1\) Five participants missed one trial due to computer malfunction and thus percentage correct was calculated from 143 rather than 144 trials. Data from these participants were not included in reliability analyses.
Adaptive expression coding task (Burton et al., 2013)

Adaptive coding ability was measured using the expression aftereffect task developed and described in detail in Burton et al. (2013). Participants adapted to weak (50%) and strong (100%) anti-expression adaptors for each of the four basic expressions used as test stimuli (happy, angry, sad, scared) (see Figure 1; Skinner & Benton, 2010 for more details). Surprised and disgusted were excluded because this task was initially developed for children, who tend to misidentify those expressions. Aftereffects are typically larger after adaptation to strong than weak adaptors (and they were in this study, \( t(87) = 7.48, p < .001, d = 1.60 \)). We were not interested in this difference but included two adaptor strengths to ensure we would obtain a good range of aftereffect strengths across participants and minimise the impact of any floor or ceiling effects.

There were 80 trials. On each trial, participants viewed an anti-expression adaptor face for a total of 15 sec (15 x 1,000 ms exposures, interspersed with 150 ms blank ISIs) and then judged the expression of a test face (shown for 400 ms after a 150 ms ISI) as one of four basic expressions (happy, angry, sad, scared). On 60% of the trials, the primary test stimulus was an average face (a morph of 100% happy, sad, angry, fearful, surprised, disgusted, and neutral expressions, Figure 1), which displays no clearly discernible basic emotion (in contrast to ‘neutral’ faces which can often resemble sadness). Following Burton et al. (2013), we calculated the effects of adaptation as the proportion of responses to average test faces that matched (i.e., were opposite to) the anti-expression of the adaptor (separately for the weaker and stronger adaptors). On the remaining trials, test stimuli displaying clearer emotions were shown (50% expressions, Figure 1). These were used to maintain motivation and ensure that participants could identify the expressions (and performance on these trials was high, \( M = .90, SD = .89 \)), but these were not used to calculate aftereffects. To minimize the effect of retinotopic adaptation, test stimuli were 75% of the size of the adaptor stimuli. One pseudo-random order of trials was created to
prevent adaptation accumulating across trials (i.e., each set of four trials contained one of each of the four anti-expressions as adaptors and no two consecutive trials could use the same anti-expression adaptor). The same order was administered to all participants so that variation in performance could not be attributed to differences in trial order between participants.

Prior to completing the adaptation task, participants were provided with training to ensure that they could identify 100%- and 50%-strength expressions using the keyboard labels (see Burton et al., for details of how the training was structured). To encourage attention to the adaptor faces, the eyes or lips of three adaptor faces in the sequence were brightened and participants were asked to press a key labelled ‘eyes’ or one labelled ‘lips’ as soon as they detected which feature changed (and participants were provided with training on how to do this before completing the task) (see Burton et al., 2013 for additional details).

**Self-report measures**

1. *The Depression Anxiety and Stress Scales (DASS) (21-item)* (Lovibond & Lovibond, 1995).

Depression and anxiety over the past week were each measured with 7-items on a 4-point scale from ‘did not apply to me at all’ to ‘applied to me very much’. The scales measure severity of symptoms across the population and have been reported as valid and reliable for non-clinical samples (Cronbach’s alpha for depression = .82; anxiety = .90; Henry & Crawford, 2005). We included the stress scale as part of the questionnaire but did not analyse this sub-scale as we had no theoretical reason to expect a relationship.


The PANAS consists of two 10-item mood scales, in which participants are asked to rate the extent to which they have experienced particular emotions during the past week on a 5-point scale from ‘very slightly or not at all’ to ‘very much’. The scales are valid and
reliable for non-clinical samples (Cronbach’s alpha for positive affect = .89; negative affect = .85; Crawford & Henry, 2004).

3. *The Autism-Spectrum Quotient (AQ) (Baron-Cohen et al., 2001).*

This widely used self-report measure of autism-like traits in adults contains 50 items, in which participants respond on ‘definitely agree’, ‘agree’, ‘disagree’ or ‘definitely disagree’. We report a total AQ score as per the original binary scoring method, in which both the ‘agree’ options are scored the same (either 0 or 1) and both the ‘disagree’ options are scored the same (either 0 or 1), for an overall score that could vary between 0 and 50.

4. *Empathy Quotient (EQ, Baron-Cohen & Wheelwright, 2004) and Basic Empathy Scale (BES, Joliffe & Farrington, 2006).*

Levels of empathy were measured with the EQ and BES. The EQ consists of 40-items that participants endorse on a 4-point scale from ‘strongly agree’ to ‘strongly disagree’ and the scale shows good validity and reliability (Cronbach’s $\alpha = .92$; Baron-Cohen & Wheelwright, 2004; Lawrence et al., 2004).

The BES contains 20 items, for which participants use a 5-point Likert scale, ranging from ‘strongly disagree’ to ‘strongly agree’. The affective (11-items) and cognitive (9-items) sub-scales demonstrate good reliability and validity (Cronbach’s $\alpha$’s for subscales above .75; Joliffe & Farrington, 2006; Dawel, Palermo, O’Kearney, & McKone, 2015) and we report the data split by these two sub-scales.

**Procedure**

The emotion-matching and emotion-labelling tasks (Palermo et al., 2013) were completed before the MERT (Bänziger et al., 2009). Participants then completed the expression aftereffect task (Burton et al. 2013). Most participants completed it in the same
test session as the emotion recognition tests, except for approximately half of the ANU participants, who completed it in a subsequent test session (which varied from 1 week to 9 months later; median = 3 months). All tasks were presented on iMac monitors and were controlled by SuperLab 4.5 (Cedrus Corp.) except for the MERT, which was administered via the web-based interface following the standard instructions\(^2\).

Questionnaires assessing depression and anxiety (DASS, Lovibond & Lovibond, 1995) and mood (PANAS, Watson et al., 1988) over the past week were completed at the beginning of the test session (note that approximately half of the ANU participants were not administered the DASS or PANAS). Questionnaires assessing empathetic and autism-like traits (EQ, Baron-Cohen & Wheelwright, 2004; BES, Joliffe & Farrington, 2006; AQ, Baron-Cohen et al., 2001) were administered at the end of testing session (n = 1 ANU participant did not complete the AQ; n = 2 UWA students did not complete the EQ).

\(^2\) Most of the participants also completed a facial expression composite effect task at the beginning of the first test session, the results of which will be reported elsewhere.
Results

Descriptive statistics, range and reliability of tasks

Table 1 reports descriptive statistics and reliability for all measures.

<p>| Table 1 |
| Descriptive Statistics and Reliability for the Tasks and Measures |</p>
<table>
<thead>
<tr>
<th>N</th>
<th>Possible Range</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>SD</th>
<th>Skew</th>
<th>Kurtosis</th>
<th>Cronbach's α</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Emotion Recognition tasks</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emotion-matching task</td>
<td>88</td>
<td>0-100% (chance = 33%)</td>
<td>47</td>
<td>89</td>
<td>73.28</td>
<td>8.19</td>
<td>-.53</td>
<td>.48</td>
</tr>
<tr>
<td>Emotion-labelling task</td>
<td>88</td>
<td>0-100% (chance = 17%)</td>
<td>62.19</td>
<td>91.67</td>
<td>77.83</td>
<td>6.04</td>
<td>-.31</td>
<td>-.06</td>
</tr>
<tr>
<td>MERT</td>
<td>86</td>
<td>0-100% (chance = 10%)</td>
<td>39.00</td>
<td>68.33</td>
<td>54.73</td>
<td>6.55</td>
<td>-.39</td>
<td>.02</td>
</tr>
<tr>
<td>Emotion Recognition PC</td>
<td>86</td>
<td>-3.29</td>
<td>1.88</td>
<td>.0</td>
<td>1.0</td>
<td>-.61</td>
<td>.61</td>
<td>-</td>
</tr>
<tr>
<td><strong>Adaptive coding</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expression aftereffect</td>
<td>88</td>
<td>0-1</td>
<td>.23</td>
<td>.94</td>
<td>.59</td>
<td>.19</td>
<td>-.11</td>
<td>1.14</td>
</tr>
<tr>
<td><strong>Self-report questionnaires</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DASS-depression</td>
<td>63</td>
<td>0-21</td>
<td>0</td>
<td>14</td>
<td>4.17</td>
<td>3.74</td>
<td>.99</td>
<td>.20</td>
</tr>
<tr>
<td>DASS-anxiety</td>
<td>63</td>
<td>0-21</td>
<td>0</td>
<td>13</td>
<td>3.05</td>
<td>3.18</td>
<td>1.31</td>
<td>1.01</td>
</tr>
<tr>
<td>PANAS-positive</td>
<td>64</td>
<td>10-50</td>
<td>10</td>
<td>50</td>
<td>29.13</td>
<td>6.06</td>
<td>.26</td>
<td>.64</td>
</tr>
<tr>
<td>PANAS-negative</td>
<td>64</td>
<td>10-50</td>
<td>10</td>
<td>31</td>
<td>17.60</td>
<td>4.96</td>
<td>.94</td>
<td>.12</td>
</tr>
<tr>
<td>AQ</td>
<td>87</td>
<td>0-50</td>
<td>4</td>
<td>34</td>
<td>16.20</td>
<td>5.78</td>
<td>.66</td>
<td>.93</td>
</tr>
<tr>
<td>EQ</td>
<td>86</td>
<td>0-80</td>
<td>18</td>
<td>68</td>
<td>44.56</td>
<td>10.61</td>
<td>-.07</td>
<td>-.24</td>
</tr>
<tr>
<td>BES-affective</td>
<td>88</td>
<td>11-55</td>
<td>18</td>
<td>54</td>
<td>40.40</td>
<td>7.26</td>
<td>-.68</td>
<td>.65</td>
</tr>
<tr>
<td>BES-cognitive</td>
<td>88</td>
<td>9-45</td>
<td>28</td>
<td>45</td>
<td>36.90</td>
<td>3.37</td>
<td>.09</td>
<td>.34</td>
</tr>
</tbody>
</table>

Notes. *Emotion Recognition Tasks*: Emotion-matching and labelling tasks (Palermo et al., 2013) (\(^a\) based on N = 83, see footnote 2); MERT = The Multimodal Emotion Recognition Test (Bänziger, et al. 2009); Emotion Recognition PC is calculated from the three emotion recognition tests (see text for details). *Expression Aftereffect*: average of both weak (50%) and strong (100%) adaptors (see text for details). *Self-report questionnaires*: DASS = The Depression Anxiety and Stress Scales (21-item) (Lovibond & Lovibond, 1995). Note that these scores can be doubled to be comparable to the full-length DASS-42 but we opted to use the raw scores to be comparable to Australian norms obtained by Crawford, Cayely, Lovibond, Wilson, & Hartley, 2011); PANAS = Positive and Negative Affect Schedule (Watson et al., 1988); EQ = Empathy Quotient (Baron-Cohen & Wheelwright, 2004); BES = Basic Empathy Scale (Joliffe & Farrington, 2006); AQ = The Autism-Spectrum Quotient (Baron-Cohen et al., 2001).
**Emotion-recognition tasks**

Percentage correct was calculated for these tasks (see Table 1). The data from two participants on the MERT were lost due to computer malfunction. The scores of one other participant were outliers on both the Emotion-labelling task \((z = -3.63, M = 54.86\%\)) and the MERT \((z = -3.38, M = 31.67\%\)), and their influence was reduced by changing each score to be one percentage point larger than the next most extreme score in the distribution (Haworth, 2006; Tabachnick & Fidell, 1989). Skew and kurtosis were within acceptable limits for parametric analysis (Table 1) (Kendall & Stuart, 1958). All tests had a good range of scores with average performance being neither at floor nor ceiling and reliability was moderately high (Table 1).

We correlated performance across the three expression tasks, and compared the correlations to the maximum possible correlation or upper bound for each correlation when accounting for measurement error (calculated as the square root of the product of the reliabilities of the two tasks; Kaplan & Saccuzzo, 1993; McKone & Palermo, 2010). All of the tasks were positively correlated \((all \ p’s < .001\): Emotion-matching and Emotion-labelling, \(r(86) = .58, r_{upperbound} = .78\); Emotion-matching and MERT, \(r(84) = .53, r_{upperbound} = .73\); Emotion-labelling and MERT, \(r(84) = .50, r_{upperbound} = .73\). The correlations are large (Gignac et al., 2016) and close to the upperbounds, indicating that the three tasks are tapping similar constructs.

Given the relatively high correlations between the three expression tasks, we also combined the scores across the three tasks using principal components analysis (PCA) to derive a measure of general ‘emotion recognition’, independent of the task. Bartlett’s test of sphericity indicated that the correlations between the three tasks were sufficient for PCA \((\chi^2 (3) = 68.21, p = .000)\) and the Kaiser-Meyer-Oklin measure also supported the use of a PCA \((KMO= .70)\). One underlying principal component was extracted that explained 69.2% of the variance in scores. We used each participant’s factor score as a measure of their
general emotion recognition ability. We also report correlations between the perceptual and affective factors and each of the three emotion-recognition tasks separately (Tables 2 and 3). We do this because the upperbounds were not identical, suggesting that the tasks are similar, but not identical, and may be tapping different processes (see Palermo et al., 2013 for discussion of similarities and differences between emotion-matching and –labelling).

Age did not correlate with emotion recognition ability, as measured by each of the three tasks or the Emotion-recognition PC, \( r \)'s < .09, \( p \)'s > .43. However, sex differences were evident, with women scoring higher than men on the Emotion-matching task, \( t(86) = 2.43, p = .02, d = 0.52, \) (Female \( M = 75.1, SD = 7.1; \) Male \( M = 70.9, SD = 9.0 \)); the Emotion-labelling task, \( t(86) = 2.33, p = .02, d = 0.49, \) (Female \( M = 79.1, SD = 5.4; \) Male \( M = 76.2, SD = 6.4 \)); and the Emotion recognition PC, \( t(84) = 2.23, p = .03, d = 0.48, \) (Female \( M = .20, SD = .93; \) Male \( M = -.27, SD = 1.03 \)), but not the MERT, \( t(84) = 2.80, p = .78, \) (Female \( M = 54.9, SD = 7.0; \) Male \( M = 54.5, SD = 5.9 \)).

**Expression aftereffect task**

Adaptation was measured as the proportion of responses to the average test face that matched the antiexpression of the adaptor (e.g., responding ‘happy’ after adapting to antihappy). A higher proportion of matching responses indicates a larger aftereffect.

Aftereffects for the weaker and stronger adaptors were strongly correlated, \( r(86) = .61, p < .001. \) For each participant we computed a score averaged across both adaptor strengths. The mean of this average aftereffect (\( M = 0.59, SD = .12 \)) was very similar to that observed by Burton et al. (2013) (\( M = 0.62, SD = 0.16 \)), and was significantly greater than chance, \( t(87) = 26.64, p < .001, d = 5.71. \) Crucially, we see a range of adaptability across this average score (Table 1). We opted to use this average score as our primary measure of adaptation strength because this score was more reliable than scores for the separate
adaptor strengths (α = .74 versus .60 for weak- and .58 strong-adaptors). Aftereffect strength did not correlate with age, r(86) < .01, p > .99 or vary by sex, t(86) = 0.65, p = .52.

Facial expression recognition ability correlates with the strength of adaptive coding of facial expression

Correlations were used to investigate whether expression recognition ability (measured with the three tasks and the Emotion-recognition PC) was linked with the strength of expression aftereffects. Given sex differences in three of the emotion recognition measures (emotion-matching, -labelling and the emotion recognition PC), we also conducted partial correlations, controlling for sex. There were no multivariate outliers, according to Mahalanobis distances. Skew and kurtosis were within acceptable limits for parametric analysis (Table 1; Stuart & Kendall, 1958).

We found significant positive correlations between expression aftereffects and expression recognition ability (Table 2, Figure 2), the size of which can be considered relatively large (Gignac et al., 2016).

Table 2.
Pearson and Partial Correlations (Controlling for Sex) Between Emotion Recognition Ability and Adaptive Expression coding

<table>
<thead>
<tr>
<th>Expression Aftereffect</th>
<th>Emotion-recognition ability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Emotion-</td>
</tr>
<tr>
<td></td>
<td>matching task</td>
</tr>
<tr>
<td>Pearson r</td>
<td>.287 (p = .007)</td>
</tr>
<tr>
<td>Partial correlation (sex)</td>
<td>.279 (p = .009)</td>
</tr>
<tr>
<td>Upper bound</td>
<td>.76</td>
</tr>
<tr>
<td>N</td>
<td>88</td>
</tr>
</tbody>
</table>

Notes. Emotion-matching and -labelling tasks (Palermo et al., 2013); MERT = The Multimodal Emotion Recognition Test (Bänziger, et al. 2009); a as there were no sex differences on the MERT (see text for details), no partial correlation is reported for the MERT; Emotion Recognition PC is calculated from the three emotion recognition tests (see text for details). Bold = p < .01.

As is evident in Figure 2, one individual’s emotion-recognition PC score (-3.289) was close to being an outlier. The correlation is maintained when this data point is excluded, r(83) = .337, p < .002.
Figure 2. Scatterplot showing the relationship between emotion recognition ability (here, the Emotion Recognition PC) and strength of facial expression aftereffects.

We also examined accuracy on the attention task performed during adaptation. Average accuracy for reporting whether the eye or lip brightness increased was high (M = .91, SD = .09), suggesting that overall the task was successful in ensuring that participants paid good attention to the adapting faces. However, there was some variability, including two individuals who performed more poorly than average (z = -4.12, M = .54; z = -3.55, M = .59) and whose scores were changed to be one percentage point larger than the next most extreme score in the distribution to reduce excessive skew and kurtosis (Haworth, 2006; Tabachnick & Fidell, 1989). Attention to the adaptors was associated with performance on the aftereffect task, $r(86) = .23$, $p = .03$. Crucially, however, the size of aftereffects remained significantly and substantially correlated with recognition ability when attention was
controlled (emotion-matching, \( r(85) = .24, p = .03 \); emotion-labelling, \( r(85) = .34, p = .001 \); MERT, \( r(85) = .28, p = .009 \); emotion recognition PC, \( r(83) = .35, p = .001 \)). Therefore, the observed relationships between adaptation and expression recognition (Table 2) cannot be attributed simply to variation in attention to faces.

**Facial expression recognition ability correlates with self-reported anxiety, but not other affective factors**

There were four univariate outliers (DASS-Depression, \( z = 3.49, \text{score} = 18 \); PANAS-negative, \( z = 4.09, \text{score} = 40 \); EQ, \( z = -3.41, \text{score} = 7 \); BES-cognitive, \( z = -3.89, \text{score} = 23 \)) and these scores were changed to be one unit larger than the next most extreme score in the distribution (Haworth, 2006; Tabachnick & Fidell, 1989). Two female participants scored 34 on the AQ, exceeding the cut-off of 32, which indicates a possible diagnosis of ASD (Baron-Cohen et al., 2001). In the absence of such a diagnosis, their scores were retained as part of the general continuum (as per Rhodes et al., 2013). There were no multivariate outliers, according to Mahalanobis distances. Skew and kurtosis were within acceptable limits for parametric analysis (Table 1; Stuart & Kendall, 1958).

Age did not correlate with any of the self-report measures except for PANAS-positive, with more positive mood associated with increasing age, \( r(62) = .26, p = .04 \) (for all other measures, \( r’s < .14, p’s > .20 \)). Sex differences were generally not apparent, except for BES-affective, \( t(86) = 4.99, p < .001, d = 1.05 \) (Female \( M = 43.4, SD = 5.4 \); Male \( M = 36.5, SD = 7.6 \)) (which is consistent with Jolliffe & Farrington, 2006) (for all other measures, \( t’s < 1.55, p’s > .09 \)).

There were significant negative correlations between self-reported levels of anxiety over the past week (measured with DASS Anxiety) and expression recognition ability, with greater anxiety linked to poorer expression recognition, although we note that these would not withstand Bonferroni correction for multiple comparisons (for 8 comparisons \( p \) would be
In contrast, the ability to recognise facial expressions was not associated with depression levels over the past week (DASS depression), or positive or negative affect over the past week (PANAS positive or negative) (Table 3). Nor were there significant correlations between expression recognition ability and levels of autism-like traits (AQ) or empathy (EQ, BES-affective or BES-cognitive) (Table 3).

---

4 The AQ can also be scored with a more recent 4-point scoring method (Hoekstra, Bartels, Cath, & Boomsma, 2008; Manera, Del Giudice, Grandi, & Colle, 2011) and can also be split into factors, such as those pertaining to more social autism-like traits versus those related to attention-to-detail (e.g., Davis et al., 2016; Manera et al., 2011; Rhodes, Jeffery, Taylor, & Ewing, 2013). We found no significant relationships with emotion recognition ability when the data were scored with the 4-point method and/or split into these two factors.
Table 3.
Pearson and Partial Correlations (Controlling for Sex, or Sex & Age) Between Emotion Recognition Ability Self-Report Questionnaires

<table>
<thead>
<tr>
<th>Self-report questionnaires</th>
<th>Emotion-recognition ability</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Emotion-matching task</td>
<td>Emotion-labeling task</td>
<td>MERT a</td>
<td>Emotion Recognition PC</td>
</tr>
<tr>
<td>DASS-Depression</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson r</td>
<td>-.041 (p = .751)</td>
<td>-.020 (p = .874)</td>
<td>.070 (p = .585)</td>
<td>.003 (p = .982)</td>
</tr>
<tr>
<td>Partial correlation (sex)</td>
<td>-.025 (p = .845)</td>
<td>.003 (p = .983)</td>
<td>--</td>
<td>.016 (p = .902)</td>
</tr>
<tr>
<td>Upper bound</td>
<td>.86</td>
<td>.86</td>
<td>.80</td>
<td>-</td>
</tr>
<tr>
<td>N</td>
<td>63</td>
<td>63</td>
<td>63</td>
<td>63</td>
</tr>
<tr>
<td>DASS-Anxiety</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson r</td>
<td>-.287 (p = .023)</td>
<td>-.255 (p = .044)</td>
<td>-.216 (p = .089)</td>
<td>-.305 (p = .015)</td>
</tr>
<tr>
<td>Partial correlation (sex)</td>
<td>-.310 (p = .014)</td>
<td>-.289 (p = .023)</td>
<td>--</td>
<td>-.324 (p = .010)</td>
</tr>
<tr>
<td>Upper bound</td>
<td>.84</td>
<td>.84</td>
<td>.80</td>
<td>-</td>
</tr>
<tr>
<td>N</td>
<td>63</td>
<td>63</td>
<td>63</td>
<td>63</td>
</tr>
<tr>
<td>PANAS-Positive</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson r</td>
<td>-.035 (p = .782)</td>
<td>-.071 (p = .576)</td>
<td>-.159 (p = .210)</td>
<td>-.106 (p = .403)</td>
</tr>
<tr>
<td>Partial correlation (sex &amp;</td>
<td>-.054 (p = .677)</td>
<td>-.056 (p = .663)</td>
<td>-.159 (p = .214)</td>
<td>-.107 (p = .407)</td>
</tr>
<tr>
<td>age)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upper bound</td>
<td>.83</td>
<td>.83</td>
<td>.78</td>
<td>-</td>
</tr>
<tr>
<td>N</td>
<td>64</td>
<td>64</td>
<td>64</td>
<td>64</td>
</tr>
<tr>
<td>PANAS -Negative</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson r</td>
<td>-.147 (p = .247)</td>
<td>-.038 (p = .765)</td>
<td>.017 (p = .893)</td>
<td>-.068 (p = .595)</td>
</tr>
<tr>
<td>Partial correlation (sex)</td>
<td>-.139 (p = .277)</td>
<td>-.026 (p = .842)</td>
<td>--</td>
<td>-.061 (p = .635)</td>
</tr>
<tr>
<td>Upper bound</td>
<td>.81</td>
<td>.81</td>
<td>.76</td>
<td>-</td>
</tr>
<tr>
<td>N</td>
<td>64</td>
<td>64</td>
<td>64</td>
<td>64</td>
</tr>
<tr>
<td>AQ</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson r</td>
<td>.035 (p = .751)</td>
<td>-.123 (p = .255)</td>
<td>-.010 (p = .927)</td>
<td>-.043 (p = .695)</td>
</tr>
<tr>
<td>Partial correlation (sex)</td>
<td>.037 (p = .732)</td>
<td>-.143 (p = .191)</td>
<td>--</td>
<td>-.044 (p = .691)</td>
</tr>
<tr>
<td>Upper bound</td>
<td>.76</td>
<td>.76</td>
<td>.71</td>
<td>-</td>
</tr>
<tr>
<td>N</td>
<td>87</td>
<td>87</td>
<td>86</td>
<td>86</td>
</tr>
<tr>
<td>EQ</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson r</td>
<td>.039 (p = .724)</td>
<td>.032 (p = .767)</td>
<td>-.004 (p = .970)</td>
<td>.033 (p = .762)</td>
</tr>
<tr>
<td>Partial correlation (sex)</td>
<td>-.001 (p = .992)</td>
<td>-.007 (p = .951)</td>
<td>--</td>
<td>-.005 (p = .966)</td>
</tr>
<tr>
<td>Upper bound</td>
<td>.83</td>
<td>.83</td>
<td>.77</td>
<td>-</td>
</tr>
<tr>
<td>N</td>
<td>86</td>
<td>86</td>
<td>84</td>
<td>84</td>
</tr>
<tr>
<td>BES-affective</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson r</td>
<td>.064 (p = .551)</td>
<td>.082 (p = .447)</td>
<td>.048 (p = .662)</td>
<td>.091 (p = .402)</td>
</tr>
<tr>
<td>Partial correlation (sex)</td>
<td>-.065 (p = .549)</td>
<td>-.039 (p = .720)</td>
<td>.038 (p = .729)</td>
<td>-.022 (p = .844)</td>
</tr>
<tr>
<td>Upper bound</td>
<td>.82</td>
<td>.82</td>
<td>.78</td>
<td>-</td>
</tr>
<tr>
<td>N</td>
<td>88</td>
<td>88</td>
<td>86</td>
<td>86</td>
</tr>
<tr>
<td>BES-cognitive</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson r</td>
<td>.157 (p = .145)</td>
<td>.178 (p = .096)</td>
<td>.104 (p = .339)</td>
<td>.195 (p = .072)</td>
</tr>
<tr>
<td>Partial correlation (sex)</td>
<td>.132 (p = .222)</td>
<td>.155 (p = .150)</td>
<td>--</td>
<td>-.177 (p = .104)</td>
</tr>
<tr>
<td>Upper bound</td>
<td>.75</td>
<td>.75</td>
<td>.70</td>
<td>-</td>
</tr>
<tr>
<td>N</td>
<td>88</td>
<td>88</td>
<td>86</td>
<td>86</td>
</tr>
</tbody>
</table>
Notes. *Self-report questionnaires*: DASS = The Depression Anxiety and Stress Scales (21-item) (Lovibond & Lovibond, 1995); PANAS = Positive and Negative Affect Schedule (Watson et al., 1988); EQ = Empathy Quotient (Baron-Cohen & Wheelwright, 2004); BES = Basic Empathy Scale (Joliffe & Farrington, 2006); AQ = The Autism-Spectrum Quotient (Baron-Cohen et al., 2001). *Emotion Recognition Tasks*: Emotion-matching and labelling tasks (Palermo et al., 2013); MERT = The Multimodal Emotion Recognition Test (Bänziger, et al. 2009); *as there were no sex differences on the MERT (see text for details), no partial correlations are reported for the MERT except for the PANAS-Positive with age and the BES-affective for sex; Emotion Recognition PC is calculated from the three emotion recognition tests (see text for details). **Bold p < .05.**

*Figure 3.* Scatterplot showing the relationship between emotion recognition ability (here, the Emotion Recognition PC) and self-reported anxiety on the DASS.
Contribution of facial expression aftereffects and self-reported anxiety to expression recognition ability

Expression aftereffect strength was not correlated with self-reported levels of anxiety, \( r(60) = .08, p = .27 \). A multiple regression was conducted to evaluate whether both adaptability to expression and anxiety were significant predictors of expression recognition ability (with the expression recognition PC as the criterion). Results of evaluation of assumptions were satisfactory after one outlier with a standardised residual > 3 was excluded.

Both of the predictors contributed significantly to expression recognition ability, \( R = .66, R^2 = .43, \) adjusted \( R^2 = .40, F(3,61) = 14.70, p < .001 \), whereas sex did not (Table 4). Thus both adaptive coding and anxiety make independent contributions to facial expression recognition ability. There was no significant difference between the standardised beta weights for expression aftereffects and anxiety (\( \Delta \beta = .16 \)), given that the absolute value of the bias-corrected confidence intervals for \( \beta \) overlap by more than 50% (Table 4) (Cumming, 2009).

Table 4.
Multiple Regression of Aftereffect Strength, Anxiety, and Sex as Predictors of Emotion Recognition Ability (as measured with the Emotion Recognition PC).

<table>
<thead>
<tr>
<th>Predictor</th>
<th>B</th>
<th>95% CI for B</th>
<th>SE</th>
<th>( \beta )</th>
<th>95% CI for ( \beta )</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expression Aftereffect</td>
<td>3.96</td>
<td>2.51 to 5.40</td>
<td>.72</td>
<td>.55</td>
<td>.32 to .68</td>
<td>.000</td>
</tr>
<tr>
<td>DASS-Anxiety</td>
<td>-0.11</td>
<td>-0.17 to -0.06</td>
<td>.03</td>
<td>-0.39</td>
<td>-0.18 to -0.61</td>
<td>.000</td>
</tr>
<tr>
<td>Sex</td>
<td>-0.15</td>
<td>-0.52 to 0.22</td>
<td>.19</td>
<td>-0.08</td>
<td>-0.26 to 0.11</td>
<td>.423</td>
</tr>
<tr>
<td>Constant</td>
<td>-1.76</td>
<td>-2.83 to -0.69</td>
<td>.53</td>
<td>-0.23</td>
<td>-0.23 to 0.15</td>
<td>.002</td>
</tr>
</tbody>
</table>

Notes. \( B \) = unstandardized beta weight; 95% CI for \( B \) = 95% confidence interval; \( SE \ B \) = standard error of \( B \); \( \beta \) = standardized beta weight; 95% CI for \( \beta \) = 95% confidence interval calculated via bias corrected bootstrapping (1,000 re-samples); \( p \) = significance value. DASS = The Depression Anxiety and Stress Scales (21-item) (Lovibond & Lovibond, 1995).
Discussion

We investigated whether the strength of a key perceptual mechanism of face perception, adaptive coding, predicts individual differences in the ability to recognise facial expressions of emotion in typical young adults. We found that it did. Stronger adaptive coding of expression was associated with better facial expression recognition. We also identified a single affective factor – anxiety – that was associated with poorer facial expression recognition. These factors made similar sized contributions but did so relatively independently, so we discuss each in turn.

Individual differences in the strength of adaptive coding of expression were associated with the ability to recognise facial expressions. That is, larger aftereffects, which indicate greater adaptability to expression, were linked with better facial expression recognition ability. This pattern of association was evident for the matching of static facial expressions, the labelling of static and dynamic facial expressions, and a PC extracted from all three tasks that measured emotion recognition ability relatively independently of the task. These findings extend the pattern of association reported in Rhodes et al. (2015) using a single expression-labelling task by showing that this positive association holds across a number of different emotion recognition tasks that have different demands, including one that does not require labelling but simply perceptual matching, as well as a task-general measure of emotion recognition ability. We also showed that this positive association cannot simply be attributed to variation in attention to faces. Overall, our results clearly support a functional role of adaptation in facial expression recognition.

One possible functional role for adaptation is that over time adaptation calibrates coding mechanisms to the population of facial expressions that we experience and this calibration facilitates recognition of subtle differences between expressions (for reviews, see Clifford & Rhodes, 2005; Rhodes & Leopold, 2011; Webster & MacLeod, 2011). Another
possibility is that adaptation and discrimination are linked because adaptation strength and discrimination sensitivity may both be modulated by the slope of tuning functions of the neural populations that code face dimensions. Steeper slopes could result in both stronger adaptation and better discrimination than shallower slopes (see Dennett et al., 2012 for a discussion).

For facial identity, evidence for the functionality of adaptive coding comes not only from individual differences studies (Rhodes et al, 2014; Rhodes et al, 2015; Engfors, Jeffery, Gignac, & Palermo, 2016), but from studies showing that aftereffects tapping identity are also reduced in populations with face recognition difficulties, such as congenital prosopagnosia (Palermo et al., 2011), autism (Pellicano et al., 2007; Ewing, Leach, Pellicano, Jeffery, & Rhodes, 2013; Ewing, Pellicano & Rhodes, 2013) and cataract removal patients (Rhodes et al, 2016). Interestingly, we have recently shown that expression aftereffects are similarly reduced in individuals with autism, a population with previously reported facial expression recognition difficulties (Rhodes, Burton, Jeffery, Read, Taylor and Ewing, Submitted).

The discussion above focusses on long-term functional benefits that might accrue from possessing a more adaptable face coding system. As noted in the Introduction, the short-term aftereffects measured in the present study are argued to sample the strength of individuals’ adaptive coding. While it is clear that adaptation may influence perceptual coding at a variety of timescales, the relationship between adaptation at different timescales is not yet certain. It is plausible, however, that short-term adaptation effects are cumulative so that, over time, they result in longer-term changes to the adapted state (e.g., the norm). We note that short-term adaptation may also have immediate functional benefits. For example, biasing perception away from the adapting expression may be beneficial for detecting changes in expression in ongoing interactions because such a bias would enhance
sensitivity to novelty (see Webster 2015 for discussion). The present study was not designed to assess such benefits though it is an interesting question for future research.

We also found that levels of self-reported anxiety over the past week (as measured with DASS anxiety) were related to expression recognition ability, in that higher levels of anxiety were associated with poorer recognition. The direction of this relationship was predicted, evident across all the emotion-recognition measures, and of a moderate to large size. Nonetheless, it worth re-stating that we used numerous affective measures in the current study and so the correlations would not be significant when corrections were applied. Given this, replication of a negative relationship between anxiety and expression recognition ability is warranted.

Impaired facial expression recognition is evident in adults diagnosed with clinical anxiety disorders, and this deficit has been suggested to reflect attentional biases, negative biases and/or emotional dysregulation (Demenescu et al., 2010). Any or all of these factors could be driving the association in the typical population, and are potential areas for further investigation. Moreover, as Demenescu and colleagues point out, expression recognition deficits may contribute to deficits in social functioning in people with clinical anxiety. It will be important for future studies to investigate whether people in the typical population who report higher levels of anxiety, which is linked with poorer facial expression recognition, also report deficits in social functioning.

An important finding was that the contribution of adaptive coding and anxiety to expression recognition were independent. Moreover, the perceptual and affective factors displayed similar sized effects. These results highlight that individual differences in expression recognition ability reflect contributions from a variety of perceptual, affective and cognitive factors. Adaptive coding may not be the only perceptual mechanism that contributes to expression recognition ability. Individual differences in face identity recognition have also been associated with the strength of holistic coding, which refers to
the integration of information across the face into a global representation (Engfors et al., 2016; Wang, Li, Fang, Tian, & Liu, 2012). Expressions are also coded holistically (Calder, Young, Keane, & Dean, 2000) and we are investigating whether the strength of this holistic coding is linked with expression recognition ability. In addition, individual features may be sufficient for the recognition of some expressions (see Tobin, Favelle, & Palermo, 2015 for discussion of features and holistic coding in expression recognition), and ability to distinguish features could also be linked with individual variation. It will be of interest to ascertain whether either of these perceptual mechanisms interact with adaptive coding and/or affective mechanisms, such as anxiety.

In contrast to the relationship with anxiety, we found no relationships between facial expression recognition and other affective factors. This included depression levels over the past week (as measured with DASS depression), or positive or negative affect over the past week (as measured with the PANAS positive and negative). These results suggest that non-clinical levels of anxiety have more of an impact on facial expression recognition than other mood and temperament measures.

We found no association between facial expression recognition ability and autism-like traits (as measured with the AQ). Similarly, Halliday et al. (2014) found no association between the AQ scores of students (N = 131) and their ability to match sequentially presented prototypical basic facial expressions. Our results, and those of Halliday and colleagues are in contrast with the results of two other recent studies. Poljac et al. (2013) compared the expression recognition ability of a group of students who scored very highly on the AQ (N = 21, M = 28.6 ± 3.9) with those scoring very low (N = 18, M = 6.3 ± 0.9). Those in the low AQ group were significantly more accurate than those with high AQ. However, this group difference was mainly apparent for subtle negative facial expressions (i.e., face morphs that were moderate intensity blends between neutral and full intensity expressions). Lewis et al. (2016) showed an association between autistic traits and the ability to label
moderate intensity facial expression blends, across a large sample (N=389). These results suggest that associations between autistic traits and facial expression recognition may only be evident with subtle expressions.

There was no relationship between facial expression recognition and levels of empathy (as measured with the EQ, BES-affective or BES-cognitive). It is possible that the facial expression tasks used here were not sufficiently sensitive to the types of processes that are linked with individual differences in empathy. Previous studies that have found positive associations between expression recognition and empathy have used tasks that involve judging social emotions (e.g., embarrassment) from just the eyes of the face (Lawrence et al., 2004), presented expressions very briefly (Besel & Yuille, 2010) or used subtle expressions formed from morphed blends between emotions (Lewis et al., 2016). By contrast, our tasks assessed perception of full-strength, basic emotions presented for longer durations. Another possibility is that understanding what another person is thinking or feeling requires more than recognition of physical aspects of the face and only variation in later cognitive and affective processes involved in understanding the meaning of expressions might be associated with empathy. Such an argument would be consistent with data from computational models showing that perception of facial structure is sufficient for facial expression recognition in the absence of any affective processing (see Calvo & Nummenmaa, 2016 for discussion; also see Olderbak & Wilhelm, 2017 for additional reasons for weak and inconsistent effects between emotion perception and empathy).

Two final issues regarding the measurement of emotion recognition deserve consideration. First, we opted to use emotion-matching and labelling tasks because these tests are sensitive to individual differences in ability. However, these commonly used types of emotion-recognition tasks may have allowed participants to use a process-of-elimination strategy (e.g., “I know it’s not these emotions so it may be this other one”). Given the potential limitations of these types of tasks, future research should consider whether other
types of tasks, such as allowing the participant to freely choose labels or rate the intensity of the emotion shown, might be valid and reliable methods of measuring individual differences in ability. If so, they would be useful additions to any battery purporting to measure general emotion-recognition ability.

Second, we investigated ‘facial expression recognition’ as a unitary ability, regardless of the specific emotion expressed by the face. While there does appear to be a common process involved in recognising all basic expressions, there may also be some specific skills used for the recognition of specific emotions. For instance, skills used to recognise happiness may be partially dissociable from those used to recognise fear (for discussion see Lewis et al., 2016; Hildebrand, Sommer, Schact, & Wilhelm, 2015; Schlegel, Grandjean, & Scherer, 2012). This difference between general and specific expression recognition may be important to consider when examining relationships with affective factors. For instance, autism-like traits may be associated with individual differences in the ability to recognise fear (c.f. Uljarevic & Hamilton, 2013 in autism) but not with general expression recognition abilities. It is worth noting, however, that one recent examination of the relationship between specific emotions and one affective factor (empathy) did not find any clear emotion-specific effects, despite a total sample of nearly 1,000 participants and using structural equation modeling (Olderbak & Wilhelm, 2017).

In addition, it seems less likely that perceptual factors, such as adaptive coding, would be expression specific. Functional explanations of adaptive coding emphasise its role in facilitating discrimination between expressions, and it is difficult to see how expression-specific variation in adaptability would contribute to expression recognition. Nevertheless, it remains an open question whether individual differences in adaptive coding are expression specific.

In conclusion, we found that individual differences in the recognition of facial expressions are independently associated with variability in adaptive coding of expression.
and self-reported anxiety. This result indicates that both perceptual and affective factors influence this crucially important skill in the general adult population.
Acknowledgments

This research was supported by the Australian Research Council (ARC) Centre of Excellence for Cognition and its Disorders (CE110001021) http://www.ccd.edu.au, an ARC Discovery Project grant to EM and RP (DP110100850), and an ARC Discovery Outstanding Researcher Award to GR (DP130102300).

Thank you to Christopher P. Benton and Andrew L. Skinner for providing the stimuli and inspiration for the expression aftereffect task.
References


Ewing, L., Leach, K., Pellicano, E., Jeffery, L., & Rhodes, G. (2013). Reduced face aftereffects in autism are not due to poor attention. *PLoS one, 8*(11), e81353. doi:10.1371/journal.pone.0081353


