**Emotion recognition ability: evidence for a supramodal factor and its links to social cognition**

Hannah L. Connolly1\*, Carmen E. Lefevre2, Andrew W. Young3, and Gary J. Lewis1

1. Department of Psychology, Royal Holloway, University of London, Egham, Surrey TW20 0EX, UK
2. Centre for Behaviour Change, Department of Clinical, Educational and Health Psychology, University College London, 1-19 Torrington Place, University College London, London, WC1E 7HB, UK
3. Department of Psychology, University of York, Heslington, York YO10 5DD, UK

\* Corresponding author email: hannah.connolly.2016@live.rhul.ac.uk

Re-submission date: 16/12/19

**Abstract**

Accurate recognition of others’ emotions is an important skill for successful social interaction. Unsurprisingly, it has been an enduring topic of interest, and notable individual differences have been observed. Despite this focus, the underlying functional architecture of this ability has not been thoroughly investigated, particularly concerning emotion recognition across different sensory domains and stimulus modalities. Using a structural equation modelling approach, Study 1 (N=284) established the structure of emotion recognition ability across three expressive domains – face, body and voice – and observed strong evidence for a superordinate ‘supramodal’ emotion recognition factor, over and above domain-specific factors. Additionally, we observed a significant moderate negative association between this superordinate factor and alexithymia. In Study 2 (N=218), findings indicated that supramodal emotion recognition ability and face identity recognition are two related but independent constructs. In Study 3 (N=249), we examined links from both supramodal emotion recognition and face identity recognition to broader cognitive ability, and observed that general intelligence was a significant predictor of supramodal emotion recognition ability. In contrast, there was no association between intelligence and face identity recognition ability. Across three independent samples, then, our findings offer strong support for an emotion recognition ability factor existing across visual and auditory domains encompassing social signals conveyed by face, body and voice, and outline its associations to broader cognitive and affective traits.

**Keywords:** emotion recognition ability; identity recognition ability; individual differences; general intelligence

**Introduction**

 The ability to recognise emotions underpins much of human social interaction by providing important information regarding the mental states and potential intentions of others. Although the extent to which expressions of emotion involve spontaneous or more deliberately manipulative social signals is debated (Crivelli & Fridlund, 2018), accurate perception of emotional expressions is associated with better social functioning (Brackett, Rivers, Shiffman, Lerner, & Salovey, 2006) and greater relationship well-being (Carton, Kessler & Pape, 1999). Conversely, impairment in emotion recognition is associated with neuropsychiatric and psychopathological disorders including (but not limited to) depression (Surguladze et al., 2004), social anxiety (Joormann & Gotlib, 2006), and borderline personality disorder (Fenske et al., 2015).

 Unsurprisingly, then, individual differences in emotion recognition ability have been a topic of broad and enduring interest. However, much of this work has been restricted to facial expressions, and this is problematic for a number of reasons. Firstly, it is clear that emotion recognition ability is not limited to expressions communicated solely by the face; affective states are expressed through additional channels, including the body and voice. Secondly, theoretical perspectives have posited that cues to emotional expression from different modalities are often closely integrated in emotion perception (Calder & Young, 2005; Schirmer & Adolphs, 2017; Young, 2018). Data from behavioural (de Gelder & Vroomen, 2000) and brain imaging studies (de Gelder, Böcker, Tuomainen, Hensen, & Vroomen, 1999; Hagan et al., 2009; Hagan, Woods, Johnson, Green, & Young, 2013; Peelen, Atkinson, & Vuilleumier, 2010) support this contention. A major driver of this type of organisation may be that cross-modal integration can be particularly useful in circumstances where signals from each separate channel are ambiguous and have significant temporal demands (Bruce & Young, 2012; de Gelder & Bertelson, 2003; Young & Bruce, 2011; Young, 2018). This is often the case in everyday life, where someone's emotional reactions may change from moment to moment, or where one channel is sometimes obscured or not available (e.g. being able to see another person's face but struggling to hear them clearly). Crucially, however, it is not yet known if individual differences in emotion recognition ability reflect this theorized integrated organisation. If, as Young (2018) has argued, emotion recognition performance reflects an ability that immediately integrates relevant cues across modalities, one should expect to see the individual differences across modalities reflecting a higher-order general latent factor. Such a factor would not offer direct evidence of an integrated system *per se*, but it would provide evidence of a coherent clustering of emotion recognition abilities, in line with recent theorising by Young (2018) and others.

Of the work to address this question to date, some support for such a model has been reported. Firstly, across two studies (NStudy 1 = 305, NStudy 2 = 295), Schlegel, Grandjean, and Scherer (2012) assessed participants using the Multimodal Emotion Recognition Test, comprising both facial and auditory stimuli for two versions of each of five basic emotions (i.e. anger, disgust, fear, happiness, and sadness). They analysed their data using confirmatory factor analysis, and noted the presence of a general emotion recognition ability factor that acted across the modalities of face and voice. More recently, Lewis, Lefevre, and Young (2016) – also using a confirmatory factor analysis approach – reported evidence for a general ability factor acting across the modalities of face and body.

The above findings are consistent with the notion of a supramodal emotion recognition ability factor. However, a more complete test of this putative architecture requires assessment of all three main communicative channels in a single sample: the above-mentioned studies addressed only two types of stimuli (faces and voices, or faces and bodies).

More generally, if emotion recognition ability is indeed supramodal in nature, a key question is whether it reflects an emotion-specific ability or instead if it is reflective of broader socio-cognitive or general cognitive abilities. With the above in mind, in the current set of studies we sought to answer the following questions: 1) does emotion recognition ability reflect a superordinate ability factor encompassing face, body, and voice?; 2) if yes, is this supramodal ability factor specific to emotion, or reflective of broader socio-cognitive or general cognitive ability?

**Study 1**

In Study 1, we addressed the possible existence of a supramodal emotion recognition ability factor by examining the factor structure of emotion recognition ability across three expressive domains – face, body, and voice. We tested five models (schematically presented in Figure 1), each representing a different theoretical perspective in the field. The first model posited uncorrelated domain-specific factors underlying recognition of emotion from the face, body and voice, in line with the perspective that distinct mechanisms underlie emotion recognition from different sources of information. The second model posited a single overarching factor acting directly on all the manifest variables, in line with the idea of a general factor underlying performance across communicative modalities. The structure of the third model included latent factors for face, body, and voice emotion recognition, but also included an emotion-general factor acting directly on all three domain-specific factors (i.e. a higher-order structure). This model was posited in line with findings that emotion recognition processes are argued to be both domain-specific and domain-general across both visual (face) and auditory (voice) modalities (e.g. Lewis et al., 2016; Schlegel et al., 2012). The fourth model posited a similar higher-order structure to the third model but with the superordinate factor encompassing only the two visual domains (faces and bodies), thus allowing voice emotion recognition to be modelled as an independent factor. The final model proposed a bifactor solution whereby a supramodal factor and three domain-specific factors all contribute direct influences on the manifest variables.

**----- Insert Figure 1 here -----**

 Additionally, we took the opportunity to examine whether this putative supramodal factor of emotion recognition ability was associated with a cluster of socio-affective traits with relevance for psychopathology: specifically, alexithymia, autism-like traits, and social anxiety. Our previous work (Lewis et al., 2016) highlighted a moderate negative association with alexithymia but no association with a short form measure of autism-like traits (the AQ-10). Here we sought to provide a confirmatory test of the link to alexithymia, and to further probe possible links with autism-like traits by assessing two core autism domains - Social Behaviour/Interactions and Attention to Detail/Numbers and Patterns by using the longer 28-item Autism Quotient (AQ-28). In addition, we sought to determine whether social anxiety is linked with emotion recognition ability. This was prompted by mixed results in this domain (e.g. Mogg, Philippot, & Bradley, 2004) alongside recent work (White, Bray, & Ollendick, 2012) highlighting that social anxiety can mimic autism-like traits and thus may be the more appropriate construct to assess in the context of understanding associations between autism-like traits and emotion recognition ability.

The studies reported here were carried out on a commonly used online data collection platform: Amazon’s Mechanical Turk, and which we have used extensively in our previous work (e.g. Lewis et al., 2016). Despite some researchers’ early concerns about online data collection, recent studies have indicated that most of the earlier preconceptions regarding web-based data collection appear unfounded, and there may be substantial advantages to be had over lab-based samples (e.g. larger, more diverse samples with high internal motivation to respond appropriately) (Gosling, Vazire, Srivastava, & John, 2004). Specifically for Amazon’s Mechanical Turk, the demographics of participants have been found to be at least as representative of the American population when compared to traditional samples (on measures of sex, age, race, and education), and are often more diverse than the traditional lab-based undergraduate student samples.

Importantly for the current paper, psychometrically sound performance in online settings has been observed in the contexts of face and voice recognition. Germine and colleagues (2012) reported comparable performance on the Cambridge Memory Face Test in several large online samples and traditional lab-based samples (Germine, Nakayama, Duchaine, Chabris, Chatterjee, & Wilmer, 2012). In a validation study of the Glasgow Voice Memory Test, Aglieri and colleagues (2017) found no significant difference between performance of an online sample (N=1120) and a controlled lab sample (N=63) (Aglieri, Watson, Pernet, Latinus, Garrido, & Belin, 2017).

**Methods**

**Participants**

A total of 308 participants (131 males) were recruited from Amazon’s Mechanical Turk service. As is the norm for online stimuli presentation, a proportion of participants experienced technical failures (e.g. videos not loading); therefore we included only those individuals who completed at least 90% (≥45 of 50) of trial blocks for each emotion and modality. In addition, we excluded participants whose responses suggested a lack of attention (repeated selection of the same response) or who showed chance level performance on at least two of the emotion recognition tasks, in line with our previous work (e.g. Connolly, Young, & Lewis, 2018). This resulted in the omission of 24 participants, producing a final sample size of 284 (117 males).

Mean age of participants was 38.1 years, and ethnicity was reported as follows: White (n=222), Hispanic (n=13), Asian (n=13), Black (n=27), Native American (n=1), Arab American (n=1), Indian American (n=3) and undisclosed (n=4). These represent a typical demographic pattern for MTurk samples (Huff & Tingley, 2015). Participants gave informed consent and ethical approval was granted by the Royal Holloway, University of London Ethics Committee.

**Emotion Recognition Measures**

 ***Face Emotion Recognition:*** To assess face emotion recognition ability, we used static images taken from the Facial Expressions of Emotion: Stimuli and Tests (FEEST) set (Young, Perrett, Calder, Sprengelmeyer, & Ekman, 2002). They comprise 10 identities each depicting five basic emotions (anger, disgust, fear, happiness, and sadness) morphed to different levels of intensity (i.e., N=50 images). The prototype images from which morphs were created came from the ‘Pictures of facial affect’ series (Ekman & Friesen, 1976). Intensity of expression was varied from neutral in 25% steps up to 125% relative to the 100% prototypical expression from the Ekman and Friesen series, using the image manipulation software Psychomorph (Tiddeman, Burt, & Perrett, 2001). These morphed stimuli were piloted in a previous study (Lewis et al., 2016), and images at varying degrees of morphed intensity were selected. This is an important step as it created appropriate levels of recognition for assessment of individual differences and in order to avoid floor/ceiling effects. This selection approach was also applied to the remaining body and voice stimulus set. The emotion recognition stimuli and task paradigm for face and body domains was the same as used in Studies 1 and 2 in Lewis et al. (2016).

 ***Body Emotion Recognition:*** To assess emotion recognition ability from the body, we used point-light displays described in Atkinson, Dittrich, Gemmell, and Young (2004). This stimulus set comprises 10 different actors portraying the same five basic emotions as the face stimuli (anger, disgust, fear, happiness, and sadness), at three levels of intensity (typical, exaggerated, very exaggerated), and lasting between 4.2 and 8 seconds in length. As with the face stimuli, Lewis et al. (2016) piloted the whole range of video stimuli, and chose 50 that displayed enough variance in recognition for individual differences purposes. These stimuli were again used in our current study.

 ***Voice Emotion Recognition:*** To assess vocal emotion recognition ability, we used stimuli drawn from the Montreal Affective Voices (MAV) set (Belin, Fillion-Bilodeau, & Gosselin, 2008), which comprises 10 different actors portraying nine different emotions (anger, disgust, fear, pain, sadness, surprise, happiness and pleasure, and neutral), each lasting between 1.45 and 2.23 seconds. We ensured this stimulus set was in line with the face and body sets of stimuli by including vocal bursts of only the five basic emotions (note, due to ceiling performance of the happiness vocal bursts, we instead used pleasure as a proxy for this emotion). Similarly to the other stimuli, the bursts were first piloted to arrive at 50 vocal stimuli (10 of each emotion) with psychometric properties comparable to the other two expressive domains.

 Each of these emotional modalities (face, body, voice) were included in the tested models as latent variables, with sum scores for each emotion as manifest variables.

**Additional Measures**

Participants were also asked to complete a series of brief questionnaires assessing the following constructs. Scale sum scores were included as manifest variables, in line with other comparable work of this kind.

 ***Autism Quotient – Short (AQ-Short):*** The AQ-Short is a 28-item instrument that assesses five domains of autism-like traits: ‘Social Skills’, ‘Routine’, ‘Switching’, ‘Imagination’, and ‘Attention to Detail/Numbers and Patterns’. The majority of confirmatory factor analysis studies in the field suggest that these domains can be reliably summarized by two higher order factors: i. Social Behaviour/Interactions; and ii. Attention to Detail/Numbers and Patterns (Hoekstra et al., 2011). The Social Behaviour factor (AQ-Social) reflects the first four domains and comprises 23 items: Social Skills (7 items), Routine (4 items), Switching (4 items) and Imagination (8 items). The Attention to Detail factor (AQ-Detail) reflects the final domain and comprises 5 items. Responses were made on a 4-point Likert scale, with 1 being ‘definitely disagree’, and 4 being ‘definitely agree’, with 14 of the 28 items being reverse-scored. Example items from the AQ-Social factor include “Reading a story, I find it difficult to work out the character’s intentions”, and examples from the AQ-Detail factor include “I am fascinated by numbers”. Scale scores were generated by summing the responses from the respective sub-scales: a higher score indicates a greater degree of autism-like traits. Cronbach’s alpha for our data was good for both sub-scales (Social Behaviour α=.81; Attention to Detail α=.78).

 ***Toronto Alexithymia Scale (TAS-20):*** The TAS-20 is a 20-item instrument that assesses alexithymic traits in three broad sub-domains: Difficulty Identifying Feelings, Difficulty Describing Feelings, and Externally Oriented Thinking (Bagby, Parker, & Taylor, 1994). Examples of items from each subscale respectively include “I am often puzzled by sensations in my body”, “People tell me to describe my feelings more”, and “I prefer to analyse problems rather than just describe them”. Responses were made on a 5-point Likert scale, with 1 being ‘strongly disagree’, and 5 being ‘strongly agree’, with five of the 20 items being reverse-scored. Total score was generated by summing the responses from all 20 items: a higher score indicates a greater degree of alexithymic traits. Scale scores for each individual sub-scale were also generated. Cronbach’s alpha for our data was acceptable-to-good for all sub-scales (Difficulty Identifying Feelings α=.89; Difficulty Describing Feelings α=.78; Externally-Oriented Thinking α=.60).

 ***Mini-Social Phobia Inventory (Mini-SPIN):*** The Mini-SPIN is a three-item measure of social anxiety disorder (Connor, Kobak, Churchill, Katzelnick, & Davidson, 2001). An example of an item is “Fear of embarrassment causes me to avoid doing things or speaking to people”. Responses were made on a 5-point Likert scale, with 1 being ‘not at all, and 5 being ‘extremely’. Total score was generated by summing the responses from all three items: higher scores indicate a greater degree of social anxiety. Cronbach’s alpha for our data was good (α=.86).

**Procedure**

Stimuli were blocked according to modality. Voice, face and body blocks were each presented to the participant once, and in the same block order. The within-block presentation order was fully randomised. In a five alternative-choice paradigm, participants had to select the emotion they perceived was being portrayed. Stimulus presentation consisted of a black screen for 500ms, a fixation cross for 750ms and a further 500ms black screen that preceded the onset of the stimulus. Face stimuli were presented for 1000ms. Body and vocal stimuli lasted for the duration of each individual video or audio clip. Participants could respond at any point following the stimulus onset. Following the three emotion recognition blocks, participants were asked to complete the self-report questionnaires. The participants were fully debriefed following completion of the survey.

**Analysis**

Structural equation models were fitted and assessed using maximum likelihood estimation with the lavaan package (Rosseel, 2012) running in R. We assessed absolute fit of our models using the Chi-Square value, the Comparative Fit Index (CFI) and the Root Mean Square Error of Approximation (RMSEA). The Chi-Square value has been the traditional measure of model fit but is particularly sensitive to large sample sizes and will often reject sensible models on this basis (Jöreskog & Sörbom, 1993). Therefore, we evaluated our proposed models with further indices. For the CFI and RMSEA, values of ≥.95 and ≤.06 respectively are viewed as indicative of good model fit (Hu & Bentler, 1999). Additionally, we also measured relative model fit using the Akaike Information Criterion (AIC: Akaike, 1973), with a lower AIC indicating better model fit.

**Results**

Descriptive statistics and correlations for study variables are shown in Table 1. Out of the 105 possible correlations, a total of 101 (81 of which in the expected positive direction) were significant at the 5% level. This pattern broadly conforms to previous work of this kind (Lewis et al., 2016).

**----- Insert Table 1 here -----**

**Confirmatory Factor Analyses**

Model fit statistics are presented in Table 2. The first model, positing three uncorrelated and independent factors, showed at best a weak fit to the data (CFI: .61, RMSEA: .09, AIC: 17344.90). The single general factor model showed a good fit to the data (CFI: .94, RMSEA: .04, AIC: 17165.73). The third model, with a higher order structure, also showed a good fit to the data (CFI: .96, RMSEA: .03, AIC: 17155.94). The model with a general visual emotion recognition factor and a specific voice emotion recognition factor, showed a poor fit to the data (CFI: .80, RMSEA: .07, AIC: 17244.37). Finally, the model positing a bifactor structure was not identified and exhibited several improper parameter estimates: specifically, the Vocal Happiness variable showed substantial negative error variance (i.e. unlikely to be due to random sampling variability), and thus indicates model misspecification. Attempts to generate a proper model solution – e.g. omitting potentially offending indicators – were not successful and so we did not probe this model any further.

**----- Insert Table 2 here -----**

It is noteworthy that the model positing a single general ability factor showed only a marginally poorer fit, and the residual variance for each of the domain-specific (i.e. face, body, and voice) ability factors was modest. With this in mind, a χ2 difference test was used as an additional point of adjudication. This analysis showed that the single factor model (with fewer estimated parameters) showed a significantly worse fit relative to the higher-order model (χ2 (3, N=284) = 15.79, *p*=.002). In sum, then, the higher-order model with a supramodal and subordinate domain-specific factors was the best fitting model and so was retained for subsequent analyses (see Figure 2).

**----- Insert Figure 2 here -----**

**Relationship between Socio-Affective Traits and Emotion Recognition Ability**

We next moved to considering the associations between the supramodal factor we identified and our socio-affective measures (autism-like traits, alexithymia, and social anxiety). We added in these variables individually to the higher-order model, allowing them to correlate with the supramodal factor.

We observed a moderate negative correlation between the supramodal emotion recognition factor and alexithymia (*r*=-.33, *p*<.001, [CI95%: -.41, -.25]). The associations with the three alexithymia sub-scales were as follows: Difficulty Identifying Feelings (*r*=-.33, *p*<.001, [CI95%: -.41, -.25]); Externally-Oriented Thinking (*r*=-.27, *p*<.001, [CI95%:

-.34, -.20]); Difficulty Describing Feelings (*r*=-.19, *p*=.006, [CI95%: -.23, -.14]).

In contrast, no significant associations emerged for either of the Autism Quotient subscales (Social Behaviour: *r*=.04, *p*=.61, [CI95%: -.09, .17]; Attention to Detail: *r*=-.12, *p*=.08, [CI95%: -.25, .01]), or the measure as a whole (*r*=.00, *p*=.97, [CI95%: -.14, .13]). Finally, the social anxiety measure showed no association with the supramodal factor (*r*=-.01, *p*=.94, [CI95%: -.14, .12]). The correlations between the emotion recognition and affective measures are shown in Table 3.

**----- Insert Table 3 here -----**

**Discussion**

The results of Study 1 provide evidence for a supramodal emotion recognition factor that operates across all three of the expressive domains, encompassing both visual (i.e. face and body) and auditory (voice) sources of information. This finding is consistent with recent theories that suggest emotion recognition abilities generalise across modalities (e.g. Young, 2018). It is also interesting that the results are consistent with cross-modal work (e.g. de Gelder & Vroomen, 2000) that indicates multisensory integration of complex social auditory and visual information, such as moving bodies, facial expressions, and verbal and nonverbal vocal bursts.

Additionally, we observed significant negative associations between the supramodal emotion recognition ability factor and broader socio-affective functioning: specifically, with alexithymia, as measured by the Toronto Alexithymia Scale. This observation is consistent with the findings reported by Lewis et al. (2016) and thus appears to be a robust association.

We did not find any association between the two sub-scales assessing autism-like traits (i.e. social behaviour/interactions difficulties; or attention to detail/ numbers and patterns) and supramodal emotion recognition ability. Similarly, no association was observed between social anxiety and supramodal emotion recognition ability. These results contrast with previous literature that reports biases in the recognition of specific emotional expressions (e.g. Joormann & Gotlib, 2006; Attwood et al., 2017). However, discrepancies may be due to testing different samples of participants (general or clinical population) or assessing different aspects of anxiety (state or trait). An alternative perspective comes from recent work suggesting that emotion recognition difficulties observed in autism may actually reflect alexithymia (Cook, Brewer, Shah, & Bird, 2013), which is often co-morbid with autism (Bird & Viding, 2014). The current findings are consistent with this position.

**Study 2**

Results from Study 1 provided evidence for the existence of a supramodal ability factor underpinning emotion recognition from the face, body, and voice. However, since this observation derives from a study involving recognition of basic emotions, it leaves open a number of questions. In Study 2, we sought to address two of these. Firstly, does this supramodal ability factor extend to naturalistic stimuli involving more complex emotional expressions? Secondly, does this supramodal factor specifically represent emotion recognition ability *per se*, or does it instead reflect broader socio-cognitive or general cognitive ability?

To examine the first question, we investigated the relationship between the supramodal ability factor and a well-used and more naturalistic test – the Reading the Mind in Films test (Golan, Baron-Cohen, Hill, & Golan, 2006) – which assesses the recognition of complex emotions and mental states from short social scenes taken from feature films.

To examine the second question, we assessed broader performance on a set of non-emotional face processing tests. For this purpose we chose the Mooney test (Mooney, 1956; Verhallen et al., 2014) as a measure of face perception, the Glasgow Face Matching Test (Burton, White & McNeill, 2010) to measure recognition of unfamiliar face identity, and an abbreviated version of the Cambridge Face Memory Test (Duchaine & Nakayama, 2006) as a measure of face learning. These tasks thus probe different aspects of face identity perception posited in widely-used cognitive and neural models (e.g. Bruce & Young, 1986; Haxby, Hoffman & Gobbini, 2000). In order to assess whether the supramodal factor was reflective of broader cognitive ability, we also assessed participants on a short measure of verbal intelligence.

**Methods**

**Participants**

A total of 252 participants (104 males) were recruited from Amazon’s Mechanical Turk service. As in Study 1, there were a number of participants who experienced technical failures, therefore we included only those individuals who completed 90% (≥13 of 15) of trial blocks for each emotion and modality. In addition, we again excluded participants whose responses suggested a lack of attention. This resulted in an omission of 34 participants, producing a final sample size of 218 (86 males).

Mean age of participants was 37.9 years, and ethnicity was reported as follows: White (n=167), Hispanic (n=14), Asian (n=10), Black (n=16), Native American (n=3), Indian American (n=3) and Other (n=5). These characteristics match the typical demographic pattern for MTurk samples (Huff & Tingley, 2015).

**Stimuli**

***Emotion Recognition Ability:*** To assess this ability, we drew on the same face, body and voice stimuli as detailed in Study 1. However, in order to reduce the time taken to complete the emotion recognition tasks and avoid any potential fatigue effects (given the additional measures used in Study 2), we selected three items from each emotion in each modality, so that each modality block comprised 15 items (5 emotions x 3 items) rather than the previous 50 (5 x 10). The selection procedure involved taking items displaying means between .50 and .80 (i.e. those items not showing floor or ceiling effects and thus suitable for individual differences research). Where a surplus of items was available, we chose those items varying in gender and age as much as possible. Where three items between this range of means were not available, we chose three that showed means closest to this range.

***Reading the Mind in Films Task (RMFT):*** To assess the first question of Study 2, we tested participants on the more naturalistic Reading the Mind in Films test (Golan et al., 2006). This test in its full form comprises 22 short video clips taken from four British feature films, and showing scenes of characters engaging in complex social interactions. The participant perceives how a particular character is feeling from a choice of four emotional categories. Following a pilot experiment on Amazon’s MTurk (n=37: 22 males), we selected only those video clips showing mean emotion recognition accuracies between .50 and .80, thus resulting in a total of 12 video clips. As before, this step is important in ensuring our stimuli sets are suitable for individual differences research.

***Mooney Face Test:*** The Mooney test was originally developed in the late 1950’s, but has recently been made suitable for online administration (Verhallen et al., 2014). The task measures perceptual closure using high contrast face images consisting of exclusively dark or light regions. Participants view an array of three images and in a three-alternative forced-choice paradigm, they are required to decide which image shows a face. In a pilot experiment involving MTurk participants (n=40: 26 males), we chose those items most suitable for individual variation (means between .70 and .88 in this case), resulting in a total of 10 items (from a total of 40 available) for the current study.

***Glasgow Face Matching Test (GFMT):*** This unfamiliar face matching test involves showing the participant pairs of photographed faces in clear frontal view with neutral expressions and taken on the same day but photographed with different cameras. The participant has to make a same/different person judgement, with unlimited stimulus presentation time. We used the shortened version of the test comprising 40 pairs (Burton et al., 2010), and since this test has been widely used in the field (Fysh & Bindemann, 2017), we deemed it unnecessary to pilot this for use in individual differences research.

***Cambridge Face Memory Test (CFMT):*** The original version of this unfamiliar face memory test was developed in order to diagnose individuals with prosopagnosia, and comprises a total of 72 items shown over three different phases (Duchaine & Nakayama, 2006). For purposes of keeping the total survey at a reasonable length, we chose to only use Phase 2 from this test. This phase consists of one image shown for 20 seconds at the start of the task in which six frontal-view target faces are viewed by the participant. The participant is then shown a series of 30 test images, each comprising a three-face array consisting of two distractor faces and one target face in a novel viewpoint. Their task is to select which one of the three faces they have learnt before, with unlimited time in which to make this decision.

***Verbal Intelligence:*** The participant views a series of words (10 in total) and must choose the answer option whose meaning is closest to the word in question. The ten items in this vocabulary test were taken from the Gallup-Thorndike Verbal Intelligence Test, Form A (Thorndike, 1942, as cited in Beaujean & Sheng, 2010) and the test has been used in large scale surveys (General Social Survey: Smith, Marsden, Hout, & Kim, 2012). It offers a reasonable proxy for general intelligence since it shows a strong correlation (.87) with measures of intelligence (Jensen, 2001).

**Procedure**

The three emotion recognition blocks were shown to the participant first, with the same fixed order (voice, face, and body) as in Study 1. The participants then completed the face recognition tasks, in the same fixed order: GFMT, CFMT, and Mooney (each with unlimited stimulus exposure time), then the RMFT, followed by the vocabulary test. The whole procedure took approximately 30 minutes. Participants were debriefed following completion.

**Analysis**

Appraisal of our models was based on the same fit criteria as used in Study 1, namely the χ2 test, the Comparative Fit Index (CFI), Root Mean Square of the Error Approximation (RMSEA), and the Akaike Information Criterion (AIC).

**Results**

Descriptive statistics and correlations between the emotion recognition variables are shown in Table 4. Out of 105 possible correlations, a total of 20 (19 of which in the expected positive direction) were significant at the 5% level. This pattern of correlations might raise concerns about the evidence for an underlying supramodal emotion recognition factor. We elected to aggregate as per our study plan, but discuss this issue in more detail below.

**----- Insert Table 4 here -----**

Correlations between the emotion recognition, face perception and verbal intelligence measures are shown in Table 5. Out of 28 possible correlations, a total of 17 were significant at the 5% level.

**----- Insert Table 5 here -----**

**Confirmatory Factor Analyses**

We first addressed whether a supramodal ability factor for emotion recognition was apparent, and particularly whether the Reading the Mind in Films Task (RMFT) showed a coherent path loading onto this general factor. A model with a latent emotion recognition ability factor loading onto face, body and voice emotion recognition and the RMFT fitted well (CFI: 1.00, RMSEA: .00, AIC: 3680.60). The RMFT showed a strong loading from the supramodal factor (.50) which was comparable to the factor loadings for the face (.53) and body emotion recognition scores (.53), and higher than the loadings of the voice scores (.28).

We next examined whether this supramodal emotion recognition factor was better understood at a more general, socio-cognitive level of abstraction, or reflected a specific ability factor. A model positing two uncorrelated latent factors – one for the emotion recognition variables and one for the non-emotion face variables – fitted the data poorly (CFI: .82, RMSEA: .07, AIC: 7233.01). A model positing a single overarching general factor fitted the data reasonably well (CFI: .92, RMSEA: .05, AIC: 7224.02). A model that specified a higher order social cognition factor loading on the emotion recognition factor and the ‘non-emotional face’ factor provided an excellent fit to the data (CFI: 1.00, RMSEA: .00, AIC: 7217.21). As in Study 1, a χ2 difference test was carried out to assess whether the higher-order model showed significant fit increment relative to the single factor model. Results suggested that the single factor model showed a significantly worse fit relative to the higher-order model (χ2 (1, N=218) = 8.82, *p*=.005).

Finally, we examined a bifactor solution; however, this model was not able to be identified and showed substantial negative error variance (i.e. unlikely to be due to random sampling variability). As such, we also examined a reduced bifactor model, including the general social cognition factor and either the emotion recognition factor (CFI: 1.00, RMSEA: .00, AIC: 7220.77) or face factor (CFI: 1.00, RMSEA: .00, AIC: 7220.00) – with both models exhibiting excellent fit. Overall, however, the higher order model was judged to be the best fitting model by the AIC and thus was taken forward for subsequent tests. This model, which has an overall social cognition factor and subordinate factors for emotion recognition (Face, Body, Voice, RMFT) and face perception (Mooney, GFMT, CFMT) is shown in detail in Figure 3. Of note, the emotion recognition factor forms an analogue of the supramodal emotion recognition factor observed in Study 1.

**----- Insert Figure 3 here -----**

**Relationship between Verbal Intelligence and Emotion Recognition Ability**

We next assessed whether verbal intelligence was predictive of the higher order general social cognition factor. A model including this parameter led to a poorly fitting model (CFI: .87, RMSEA: .06, AIC: 8062.05). We were concerned that this poor fit reflected the high verbal requirements of the RMFT, and the fact that the omission of a pathway between verbal intelligence and the RMFT essentially models these constructs as wholly unrelated; as such an additional loading directly from verbal intelligence to the RMFT was included. This model provided an excellent fit (CFI: .98, RMSEA: .02, AIC: 8048.26). Verbal intelligence was a modest-to-moderate predictor of the general factor (.40), and a modest predictor of the RMFT (.30). The final model is detailed in Figure 4, and the fit statistics for all models are shown in Table 6.

**----- Insert Figure 4 here -----**

**----- Insert Table 6 here -----**

**Discussion**

The results of Study 2 provide further support for the existence of a supramodal emotion recognition ability factor, and suggest that this supramodal factor extends beyond artificial laboratory-based stimuli to include more naturalistic emotional stimuli, as assessed by the Reading the Mind in Films Test (RFMT). In addition, results suggest the existence of a 'non-emotional' face perception ability factor. Most notably, our results indicate that these two latent factors represent highly related abilities – consistent with the existence of a relatively general overall socio-cognitive ability factor. Finally, we found that verbal intelligence was moderately associated with the general social cognition ability factor.

We note that the magnitude of the correlations between the individual emotion variables in Study 2 were modest and in several cases not significantly different from zero. This of course suggests that our necessarily brief measure of supramodal emotion recognition ability was noisier than desirable. Nonetheless, we still observed significant correlations across the emotion domains, indicating that these individual items did not merely capture noise, and that as well as being part of our study plan, our aggregation was reasonably principled.

**Study 3**

While the findings of Study 2 provide fairly clear evidence for the existence of a socio-cognitive ability factor, the brief and low-fidelity verbal intelligence measure did not allow strong assertions concerning the degree of association between socio-cognitive ability and general intelligence. With this mind, in Study 3 we sought to further probe the relationship between socio-cognitive ability and general intelligence with a broader test of intelligence that included both numerical and verbal components. Despite noting some reservations with our brief emotion recognition measure from Study 2, we elected to use the same assessment. This decision was largely a pragmatic one: including the longer emotion recognition ability measure alongside our other measures would have made our battery too long for maintaining high quality data collection. In addition, we wanted to further examine the quality of these items so as to determine if they represented a viable means to assess supramodal emotion recognition ability.

**Methods**

**Participants**

A total of 283 participants (147 males) were recruited from Amazon’s Mechanical Turk service. There was again a proportion of participants who experienced technical failures, therefore we included only those individuals who completed 90% (≥13 of 15) of trial blocks for each emotion and modality. In addition, we excluded participants whose responses indicated a lack of attention. This resulted in an omission of 34 participants, producing a final sample size of 249 (126 males).

Mean age of participants was 35.9 years, and ethnicity was reported as follows: White (n=185), Hispanic (n=12), East Asian (n=12), Black (n=26), Native American (n=2), Indian American (n=8), Middle Eastern (n=1) and Other (n=3). These participants represent the typical demographic pattern for MTurk samples (Huff & Tingley, 2015).

**Stimuli and Measures**

***Alice Heim 4 Test of General Intelligence (AH4) Part 1 (Heim, 1970):*** This test consists of 65 items which are either numerical or verbal in nature. Half of the questions require multiple choice responses (from a choice of 5), and half require open entry ‘creative answers’. The participants have a maximum time limit of 10 minutes to complete as many of the items as they can, after which the page automatically advances to the next section of the survey. They are encouraged to complete the questions in the order in which they are given, but are also told they may skip to the next questions if they become stuck. The participants are awarded one point for every correct answer, and their total score is generated out of a maximum of 65.

***Emotion Recognition Ability:*** To assess this ability, we used the same abbreviated face, body and voice batteries of stimuli as detailed in Study 2. As before, each modality block comprised 15 items (3 items for each of the 5 basic emotions).

***Facial Recognition Tasks*:** The latent 'non-emotional' face factor in Study 2 showed the weakest loading onto the Mooney score, and the highest loadings onto the Glasgow Face Matching and Cambridge Face Memory tasks. Therefore we chose to use only these two latter tasks.

***Glasgow Face Matching Test (GFMT):*** This test is fully outlined in Study 2. For time-saving purposes, here we generated an abbreviated version that was two-thirds of the length of the original Short version. Specifically, we took the 27 best performing items that showed no evidence of ceiling or floor effects and thus were suitable for individual differences research (i.e. these had means between .66 and .87).

***Cambridge Face Memory Test (CFMT):*** This test is fully outlined in Study 2. Again, for time-saving purposes we shortened this phase to two-thirds of its original length, by including only the 20 items with accuracy means most suitable for individual differences research (i.e. these had means between .40 and .62).

**Procedure**

The general intelligence test was shown to the participants first. They had a maximum of 10 minutes in which to answer as many questions as possible. Participants then completed the three emotion recognition blocks, with the same fixed order (voice, face, and body) as in Studies 1 and 2. This was followed by the facial recognition tasks, in the same fixed order: GFMT and CFMT. The whole procedure took approximately 25 minutes. Participants were debriefed following completion.

**Analysis**

Appraisal of our models was based on the same fit criteria as used in Studies 1 and 2, namely the χ2 test, the Comparative Fit Index (CFI), Root Mean Square of the Error Approximation (RMSEA), and the Akaike Information Criterion (AIC). Latent factors typically require three or more indicators to ensure stable identification (Floyd & Widaman, 1995). To this end, we used the first and second half of the GFMT and CFMT such that the face perception factor loaded onto four indicators. We also allowed residual correlations between the two halves of each test, as one would expect a test specific association between them. Importantly, if these associations were not formally modelled, it could lead to an unwarranted rejection of the whole model.

**Results**

Descriptive statistics and correlations between emotion variables are shown in Table 7. Out of 105 possible correlations, a total of 52 (51 of which in the expected positive direction) were significant at the 5% level. Of note, we observed a pattern of correlations between the emotion variables that more clearly supported the aggregation of these variables, although the correlations were still typically modest. This is discussed further below.

**----- Insert Table 7 here -----**

Correlations between the emotion recognition, face perception and general intelligence measures are shown in Table 8. Out of 15 possible correlations, a total of 11 were significant at the 5% level.

**----- Insert Table 8 here -----**

**Confirmatory Factor Analyses**

We first sought to assess the fit of the models that were tested in Study 2. A model positing uncorrelated emotion recognition and 'non-emotional' face perception factors was not identified. A model positing a single overarching factor also provided a poor fit to the data (CFI: .63, RMSEA: .21, AIC: 7753.85). We next examined a model that specified a higher order social cognition factor loading on the emotion recognition factor and the face factor. This model provided a good fit to the data (CFI: .97, RMSEA: .06, AIC: 7617.56), and is shown in Figure 5. Of note, this structure mirrors that of the higher-order model in Study 2 (Figure 3), suggesting that it is robust across independent samples.

**----- Insert Figure 5 here -----**

Finally, for purposes of completion, we assessed a bifactor model with the general social cognition factor acting directly on the manifest variables, but as in the previous two studies, this was not able to be identified. A reduced bifactor model with a general socio-cognitive factor and a specific emotion recognition factor showed an excellent fit (CFI: 1.00, RMSEA: .00, AIC: 7606.81. In contrast, a bifactor model with a general socio-cognitive factor and specific face perception factor was not identified.

The reduced bifactor model showed a better fit compared to the higher order model, although the difference was relatively small. In line with the higher-order model showing the best fit in Study 2, we retained this model for testing the association to general intelligence, and we report the reduced bifactor model in the Supplementary Materials (Figure S1).

**Relationship between General Intelligence and Emotion Recognition Ability**

In the higher order model, we observed a path from general intelligence to the general socio-cognitive factor that was large in magnitude (.59); however, this model showed a relatively poor fit to the data (CFI: .91, RMSEA: .10, AIC: 9571.33). Accordingly, we explored whether general intelligence instead showed direct links to either the emotion or face factors. A model with general intelligence directly predicting the emotion recognition factor (with no path to the general socio-cognitive factor) showed a good fit to the data (CFI: .98, RMSEA: .05, AIC: 9539.48), and the pathway from general intelligence to emotion recognition was significant and substantial in magnitude (.61). In contrast, a model with general intelligence directly predicting face perception (again, with no path to the general socio-cognitive factor) did not show a good fit to the data (CFI: .82, RMSEA: .14, AIC: 9614.27), and the pathway from general intelligence to face perception was modest (-.12) and non-significant.

As an additional test, we examined a model with general intelligence directly loading on both emotion recognition and the general socio-cognitive factor. This model also fitted the data well (CFI: .98, RMSEA: .05, AIC: 9539.64). Of note, the path from general intelligence to emotion recognition was still significant and substantial in magnitude (.55), whereas the path from general intelligence to the general socio-cognitive factor was modest in magnitude (.15) and non-significant. This model is detailed in Figure 6.

**----- Insert Figure 6 here -----**

We next performed the equivalent tests with the reduced bifactor specification. Including a direct path from general intelligence to the general socio-cognitive factor produced an unidentified model. We next examined whether a model including an additional path to the emotion recognition factor was identified. This model was identified and fitted the data well (CFI: 1.00, RMSEA: .00, AIC: 9527.17). The path from general intelligence to the emotion recognition factor was significant and substantial in magnitude (.66), whereas the path from general intelligence to the socio-cognitive factor was modest in magnitude (.12) and non-significant. This model is detailed in the Supplementary Materials (Figure S2). The fit values for all tested models are shown in Table 9.

**----- Insert Table 9 here -----**

**Subsidiary Analysis of Data from Study 2**

Given these findings from Study 3 of the emotion recognition factor being significantly predicted by general intelligence, we took an exploratory step of reanalysing data from Study 2 to assess whether verbal intelligence in Study 2 was more related to either the emotion recognition factor or the Reading the Mind in Films Test (RMFT) score. In a hierarchical model with verbal intelligence predicting the higher order socio-cognitive factor and the RMFT, we also included an additional path from verbal intelligence to emotion recognition, and this model showed an excellent fit to the data (CFI: .99, RMSEA: .02, AIC: 8048.32). Verbal intelligence was a significant and moderate predictor of both the socio-cognitive factor (.27) and the RMFT (.28), but showed no significant association with the emotion recognition factor. This suggests that verbal intelligence does not predict emotion recognition ability, once the association between the RMFT and broader socio-cognitive perceptual ability has been accounted for.

**Discussion**

In line with the previous two studies, the results of Study 3 offer support for the existence of a distinct supramodal emotion recognition ability factor across face, body and voice stimuli. Consistent with Study 2, a model positing two distinct, albeit highly related, factors reflecting emotion recognition and 'non-emotional' face perception abilities with a general social cognition factor in a hierarchical structure fitted the data better than a single overarching general factor model. Secondly, we noted a strong positive association between supramodal emotion recognition ability and general intelligence, broadly in line with previous studies in the field (Lewis et al., 2016; Connolly et al., 2018).

In contrast, no significant association was observed between general intelligence and either the face perception or the general socio-cognitive factors. This finding contrasts with earlier work showing an association between fluid intelligence and unfamiliar face identity matching, as measured by the Benton Facial Recognition Test (Connolly et al., 2018), but is consistent with other studies that report no correlation between intelligence and face learning, as measured by the Cambridge Face Memory Test (Wilmer et al., 2010; Palermo, O’Connor, Davis, Irons, & McKone, 2013). These studies posit that *perception* of face identity is a specific cognitive ability separate from general intelligence, and our findings here support this contention.

**General Discussion**

Being able to accurately interpret others’ emotional expressions is an important skill for social interaction. However, to date, there has been little knowledge concerning whether individual differences in this ability reflect domain-specific or superordinate processes. In three studies, we sought to address this issue by examining the structure of individual differences in emotion recognition ability across the face, body, and voice, and the associations of this ability to various affective and cognitive traits.

The findings of these studies offer several interesting observations. Firstly, in Study 1, we provide strong support for a supramodal emotion recognition ability factor that extends beyond visual domains (facial and bodily stimuli) to also include recognition of auditory emotional stimuli from the voice. Notably, evidence for the existence of this factor was also apparent in the findings from Studies 2 and 3 and therefore appears to represent a highly replicable pattern. Secondly, we observed a moderate negative association between the supramodal emotion recognition factor and alexithymia. In contrast, no significant associations with either autism-like traits or social anxiety were seen.

In Study 2, results showed that the supramodal emotion recognition factor is also linked to recognition of complex emotional stimuli that are arguably more representative of everyday scenarios. Nonetheless, the emotion recognition factor was found to be distinct from, albeit moderately-to-strongly related to, a relatively 'non-emotional' face factor derived from various tests of face perception. Emergence of this factor is consistent with recent work reporting a general face perception factor ‘*f*’, distinct from other cognitive abilities(Verhallen, et al., 2017; McCaffery, Robertson, Young & Burton, 2018). These authors’ work assessed identity perception and other non-emotional face tasks, and therefore our observations both support the proposed existence of the *f* factor and expand its scope. Our data also suggest that *f* is related to but still somewhat distinct from other high-level visual perception abilities including emotion recognition, even when the emotional stimuli comprise facial expressions. A third latent factor that encompassed the commonality between emotion and face recognition factors in a higher order structure was posited to reflect social cognition, and this factor showed a moderate association with a brief measure of verbal intelligence.

In Study 3, results also suggested the existence of two distinct latent factors: one capturing supramodal emotion recognition ability and the other capturing face perception. As in Study 2, the commonality between these two factors was considerable and was modelled as a higher order factor that we again suggest is reflective of general social cognition. In addition, the results of Study 3 showed that general intelligence was a strong predictor of the emotion recognition factor. Accurate recognition of emotion thus appears to reflect broader, non-affective cognitive processes, and this may be for a number of reasons. In particular, our emotion recognition tasks used a multiple-choice paradigm with examples of basic emotional expressions that were selected on the basis that they were sufficiently difficult to avoid ceiling performance. As such, when participants were unsure as to the target emotion, some may have resorted to cognitive strategies to ‘solve’ the intended expression, such as by a process of elimination. Implementation and success of these strategies may reflect the individual’s general cognitive ability. The tasks may also demand holding vocal or dynamic visual information in mind whilst simultaneously attending to five emotion labels in order to make a choice, and these working memory demands may conceivably underlie the strong relationship between emotion recognition and our test of general intelligence (Palermo et al., 2013). However, it should be noted that whilst emotion recognition and intelligence are related, they do not reflect the same construct. We note also that the presentation of the emotion stimuli lasted for the duration of each individual item (1 second for facial stimuli, between 4.2 and 8 seconds for body stimuli, and between 1.45 and 2.23 seconds for voice stimuli), and that participants had unlimited time in which to give their response. Therefore, the tests were likely capturing their ability to accurately perceive emotion, rather than simply a measure of processing speed.

In contrast to emotion recognition, no significant associations were observed between general intelligence and the general social cognition or non-emotional face perception factors. Previous research has noted the distinction between general intelligence and face perception (Wilmer et al., 2010), so the absence of association between these factors is not surprising. The face identity tests included here presented items with no time limit, and participants had as long as they needed to decide on their response. Therefore, as before, it is unlikely that processing speed or reaction time was being captured in performance on these tests.

 Moreover, this further supports Verhallen et al.’s (2017) findings that neither the putative *f*’ factor nor any of their individual face recognition measures show an association with scores on the standard British school qualification: *General Certificate of Secondary Education* (GCSE). In Study 3, we likewise observed no significant relationship between general intelligence and our latent face perception factor or either of the two face recognition tasks, in line with the idea that face recognition ability may be distinct from broader cognition. Our findings extend the pattern found between Verhallen’s largely undergraduate-based sample and school test results to the general intelligence scores we measured in a more diverse sample of adults at the same time point as the face tasks, and suggest this independence of abilities is a robust finding and is stable over time.

Some possibilities for future research are noteworthy. Firstly, our online sample consisted solely of US participants, and so our results may not generalize to a non-Western sample. Studies of the kind reported here in broader samples would be of value. Secondly, as highlighted in Study 2, the inter-correlations of items on the shorter emotion recognition measure were limited, and as such, we advise some caution in using abbreviated scales of this kind. The limitations of abbreviated measures are a challenge for many fields of social psychology, including personality, where short form measures are commonly used, despite concerns over their validity (Bakker & Lelkes, 2018). There are clearly situations in which it is important to keep the battery length reasonable so as to avoid participant fatigue; however, the brevity of a shorter measure will often result in lower fidelity and construct validity. The balancing of test length and validity, as well as considering power, will be an ongoing challenge, and the wider community should bear this in mind when deciding which measures to include in their future studies.

**Conclusion**

In summary, across three independent samples, our findings provide support for a robust supramodal emotion recognition ability factor that underlies the recognition of expressions across face, body and voice, and that is in itself somewhat distinct from broader face identity perception and recognition. In addition, we have outlined the association of the supramodal emotion recognition factor to various important affective and cognitive traits, specifically alexithymia and general intelligence, as well as demonstrating that it can extend to recognising more complex and naturalistic emotional stimuli.

**Supplementary Material**

These data are available on the Open Science Framework <https://osf.io/bkgt7/>

 **References**

Aglieri, V., Watson, R., Pernet, C., Latinus, M., Garrido, L., & Belin, P. (2017). The Glasgow Voice Memory Test: Assessing the ability to memorize and recognize unfamiliar voices. *Behavior Research Methods, 49*(1), 97-110.

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. *Second International Symposium on Information Theory, (1973)*, 267–281.

Atkinson, A. P., Dittrich, W. H., Gemmell, A. J., & Young, A. W. (2004). Emotion Perception from Dynamic and Static Body Expressions in Point-Light and Full-Light Displays. *Perception*, *33*(6), 717–746. <https://doi.org/10.1068/p5096>

Attwood, A. S., Easey, K. E., Dalili, M. N., Skinner, A. L., Woods, A., Crick, L., … Munafò, M. R. (2017). State anxiety and emotional face recognition in healthy volunteers. *Open Science*, *4*(5), 160855. <https://doi.org/10.1098/rsos.160855>

Bagby, R. M., Parker, J. D. A., & Taylor, G. J. (1994). The twenty-item Toronto Alexithymia scale—I. Item selection and cross-validation of the factor structure. *Journal of Psychosomatic Research*, *38*(1), 23–32. [https://doi.org/10.1016/0022-3999(94)90005-1](https://doi.org/10.1016/0022-3999%2894%2990005-1)

Beaujean, A. A., & Sheng, Y. (2010). Examining the Flynn effect in the General Social Survey Vocabulary test using item response theory. *Personality and Individual Differences, 48*(3), 294-298.

Belin, P., Fillion-Bilodeau, S., & Gosselin, F. (2008). The Montreal Affective Voices: A validated set of nonverbal affect bursts for research on auditory affective processing. *Behavior Research Methods*, *40*(2), 531–539. <https://doi.org/10.3758/BRM.40.2.531>

Bird, G., & Viding, E. (2014). The self to other model of empathy: providing a new framework for understanding empathy impairments in psychopathy, autism, and alexithymia. *Neuroscience & Biobehavioral Reviews, 47*, 520-532.

Brackett, M. A., Rivers, S. E., Shiffman, S., Lerner, N., & Salovey, P. (2006). Relating emotional abilities to social functioning: a comparison of self-report and performance measures of emotional intelligence. *Journal of Personality and Social Psychology, 91*(4), 780-795.

Bruce, V., & Young, A. W. (1986). Understanding face recognition. *British Journal of Psychology, 77*, 305-327.

Bruce, V, & Young, A. W. (2012). *Face perception*. Hove, East Sussex: Psychology Press.

Burton, A. M., White, D., & McNeill, A. (2010). The Glasgow face matching test. *Behavior Research Methods, 42*(1), 286-291. <https://doi:10.3758/BRM.42.1.286>

Calder, A. J., & Young, A. W. (2005). Understanding the recognition of facial identity and facial expression. *Nature Reviews Neuroscience*, *6*(8), 641–651. <https://doi.org/10.1038/nrn1724>

Carton, J. S., Kessler, E. A., & Pape, C. L. (1999). Nonverbal Decoding Skills and Relationship Well-Being in Adults. *Journal of Nonverbal Behavior, 23*(1), 91-100.

Connolly, H. L., Young, A. W., & Lewis, G. J. (2019). Recognition of facial expression and identity in part reflects a common ability, independent of general intelligence and visual short-term memory. *Cognition and Emotion*, *33*(6), 1119-1128.

Connor, K. M., Kobak, K. A., Churchill, L. E., Katzelnick, D., & Davidson, J. R. T. (2001). Mini‐SPIN: A brief screening assessment for generalized social anxiety disorder. *Depression and Anxiety*, *14*(2), 137–140. <https://doi.org/10.1002/da.1055>

Cook, R., Brewer, R., Shah, P., & Bird, G. (2013). Alexithymia, not autism, predicts poor recognition of emotional facial expressions. *Psychological Science, 24*(5), 723-732.

Crivelli, C., & Fridlund, A. J. (2018). Facial displays are tools for social influence. *Trends in Cognitive Sciences, 22*, 388-398.

de Gelder, B., & Bertelson, P. (2003). Multisensory integration, perception and ecological validity. *Trends in Cognitive Sciences*, *7*(10), 460–467. <https://doi.org/10.1016/j.tics.2003.08.014>

de Gelder, B., Böcker, K. B. E., Tuomainen, J., Hensen, M., & Vroomen, J. (1999). The combined perception of emotion from voice and face: early interaction revealed by human electric brain responses. *Neuroscience Letters*, *260*(2), 133–136. [https://doi.org/10.1016/S0304-3940(98)00963-X](https://doi.org/10.1016/S0304-3940%2898%2900963-X)

de Gelder, B., & Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cognition & Emotion*, *14*(3), 289–311. <https://doi.org/10.1080/026999300378824>.

Duchaine, B., & Nakayama, K. (2006). The Cambridge Face Memory Test: Results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. *Neuropsychologia, 44*(4), 576-585. https://doi: 10.1016/j.neuropsychologia.2005.07.001

Ekman, P., & Friesen, W. V. (1976). *Pictures of facial affect.* Consulting Psychologists Press.

Fenske, S., Lis, S., Liebke, L., Niedtfeld, I., Kirsch, P., & Mier, D. (2015). Emotion recognition in borderline personality disorder: effects of emotional information on negative bias. *Borderline Personality Disorder and Emotion Dysregulation*, *2*, 10. <https://doi.org/10.1186/s40479-015-0031-z>

Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological assessment, 7*(3), 286-299.

Fysh, M. C., & Bindemann, M. (2017). The Kent Face Matching Test. *British Journal of Psychology*, *109*(2), 219–231. <https://doi.org/10.1111/bjop.12260>

Germine, L., Nakayama, K., Duchaine, B. C., Chabris, C. F., Chatterjee, G., & Wilmer, J. B. (2012). Is the Web as good as the lab? Comparable performance from Web and lab in cognitive/perceptual experiments. *Psychonomic Bulletin & Review*, *19*(5), 847-857.

Golan, O., Baron-Cohen, S., Hill, J. J., & Golan, Y. (2006). The “Reading the Mind in Films” Task: Complex emotion recognition in adults with and without autism spectrum conditions. *Social Neuroscience*, *1*(2), 111–123. <https://doi.org/10.1080/17470910600980986>

Hagan, C. C., Woods, W., Johnson, S., Calder, A. J., Green, G. G. R., & Young, A. W. (2009). MEG demonstrates a supra-additive response to facial and vocal emotion in the right superior temporal sulcus. *Proceedings of the National Academy of Sciences*, *106*(47), 20010–20015. <https://doi.org/10.1073/pnas.0905792106>

Hagan, C. C., Woods, W., Johnson, S., Green, G. G. R., & Young, A. W. (2013). Involvement of right STS in audio-visual integration for affective speech demonstrated using MEG. *PLoS One, 8,* e70648.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences, 4*, 223-233.

Heim, A. W. (1970). *AH 4 group test of general intelligence manual.* NFER Publishing Company Limited.

Hoekstra, R. A., Vinkhuyzen, A. A. E., Wheelwright, S., Bartels, M., Boomsma, D. I., Baron-Cohen, S., … Sluis, S. van der. (2011). The Construction and Validation of an Abridged Version of the Autism-Spectrum Quotient (AQ-Short). *Journal of Autism and Developmental Disorders*, *41*(5), 589–596. <https://doi.org/10.1007/s10803-010-1073-0>

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. <https://doi.org/10.1080/10705519909540118>

Huff, C., & Tingley, D. (2015). “Who are these people?” Evaluating the demographic characteristics and political preferences of MTurk survey respondents. *Research & Politics*, *2*(3), 2053168015604648. <https://doi.org/10.1177/2053168015604648>

Jensen, A. (2001). Vocabulary and general intelligence. *Behavioral and Brain Sciences, 24*(6), 1109-1110. doi:10.1017/S0140525X01280133

Joormann, J., & Gotlib, I. H. (2006). Is this happiness I see? Biases in the identification of emotional facial expressions in depression and social phobia. *Journal of Abnormal Psychology*, *115*(4), 705–714. <https://doi.org/10.1037/0021-843X.115.4.705>

Jöreskog, K. G., & Sörbom, D. (1993). *LISREL 8: Structural Equation Modeling with the SIMPLIS Command Language*. Scientific Software International.

Lewis, G. J., Lefevre, C. E., & Young, A. W. (2016). Functional architecture of visual emotion recognition ability: A latent variable approach. *Journal of Experimental Psychology: General*, *145*(5), 589–602. <https://doi.org/10.1037/xge0000160>

McCaffery, J., Robertson, D. I., Young, A. W., & Burton, A. M. (2018). Individual differences in face identity processing. *Cognitive Research: Principles and Implications, 3*(1), 21. <https://doi.org/10.1186/s41235-018-0112-9>

Mogg, K., Philippot, P., & Bradley, B. P. (2004). Selective attention to angry faces in clinical social phobia. Journal of abnormal psychology, 113(1), 160-165.

Mooney, C. M. (1956). Closure with negative after-images under flickering light. *Canadian Journal of Psychology, 10*(4), 191–199.

Palermo, R., O’Connor, K. B., Davis, J. M., Irons, J., & McKone, E. (2013). New tests to measure individual differences in matching and labelling facial expressions of emotion, and their association with ability to recognise vocal emotions and facial identity. *PloS one, 8*(6), e68126.

Peelen, M. V., Atkinson, A. P., & Vuilleumier, P. (2010). Supramodal Representations of Perceived Emotions in the Human Brain. *Journal of Neuroscience*, *30*(30), 10127–10134. <https://doi.org/10.1523/JNEUROSCI.2161-10.2010>

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling and more. Version 0.5–12 (BETA). *Journal of Statistical Software*, *48*(2), 1-36.

Schirmer, A., & Adolphs, R. (2017). Emotion Perception from Face, Voice, and Touch: Comparisons and Convergence. *Trends in Cognitive Sciences*, *21*(3), 216–228. <https://doi.org/10.1016/j.tics.2017.01.001>

Schlegel, K., Grandjean, D., & Scherer, K. R. (2012). Emotion recognition: Unidimensional ability or a set of modality- and emotion-specific skills? *Personality and Individual Differences*, *53*(1), 16–21. <https://doi.org/10.1016/j.paid.2012.01.026>

Smith, T., Marsden, P. V., Hout, M., & Kim, J. (2012). *General Social Surveys,*

*1972–2012 Cumulative Codebook.* Chicago: National Opinion Research Centre.

Surguladze, S. A., Young, A. W., Senior, C., Brébion, G., Travis, M. J., & Phillips, M. L. (2004). Recognition accuracy and response bias to happy and sad facial expressions in patients with major depression. *Neuropsychology, 18*(2), 212.

Tiddeman, B., Burt, M., & Perrett, D. (2001). Prototyping and transforming facial textures for perception research. *IEEE Computer Graphics and Applications*, *21*(4), 42–50. <https://doi.org/10.1109/38.946630>

Verhallen, R. J., Bosten, J. M., Goodbourn, P. T., Bargary, G., Lawrance-Owen, A. J., & Mollon, J. D. (2014). An online version of the Mooney Face Test: phenotypic and genetic associations*. Neuropsychologia, 63*, 19-25. <http://dx.doi.org/10.1016/j.neuropsychologia.2014.08.011>

Verhallen, R. J., Bosten, J. M., Goodbourn, P. T., Lawrance-Owen, A. J., Bargary, G., & Mollon, J. D. (2017). General and specific factors in the processing of faces. *Vision Research, 141*, 217-227.

White, S. W., Bray, B. C., & Ollendick, T. H. (2012). Examining Shared and Unique Aspects of Social Anxiety Disorder and Autism Spectrum Disorder Using Factor Analysis. *Journal of Autism and Developmental Disorders*, *42*(5), 874–884. <https://doi.org/10.1007/s10803-011-1325-7>

Wilmer, J. B., Germine, L., Chabris, C. F., Chatterjee, G., Williams, M., Loken, E., … Duchaine, B. (2010). Human face recognition ability is specific and highly heritable. *Proceedings of the National Academy of Sciences*, *107*(11), 5238–5241. <https://doi.org/10.1073/pnas.0913053107>

Young, A. W. (2018). Faces, people and the brain: the 45th Sir Frederic Bartlett Lecture. *Quarterly Journal of Experimental Psychology, 71*, 569-594.

Young, A. W., & Bruce, V. (2011). Understanding person perception. *British Journal of Psychology, 102*, 959-974.

Young, A., Perrett, D., Calder, A., Sprengelmeyer, R., & Ekman, P. (2002). *Facial expressions of emotion – stimuli and tests (FEEST).* Thames Valley Test Company, Bury St Edmunds, England.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Mean | SD | 1.1 | 1.2 | 1.3 | 1.4 | 1.5 | 2.1 | 2.2 | 2.3 | 2.4 | 2.5 | 3.1 | 3.2 | 3.3 | 3.4 |
| Faces |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 1.1 Anger | .50 | .21 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 1.2 Disgust | .59 | .21 | .18\*\* [.07, .30] |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 1.3 Fear | .67 | .18 | **.31** [.20, .41] | .20\*\* [.09, .31] |  |  |  |  |  |  |  |  |  |  |  |  |
| 1.4 Happiness | .81 | .16 | **.27** [.16, .37] | .18\*\* [.07, .30] | .17\*\* [.05, .28] |  |  |  |  |  |  |  |  |  |  |  |
| 1.5 Sadness | .43 | .21 | **.31** [.20, .41] | .16\*\* [.04, .27] | **.24** [.13, .35] | **.25**[.14, .36] |  |  |  |  |  |  |  |  |  |  |
| Bodies |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 2.1 Anger | .67 | .22 | **.37**[.27, .47] | .10[-.02, .21] | **.26**[.14, .36] | **.31**[.20, .41] | **.34**[.23, .44] |  |  |  |  |  |  |  |  |  |
| 2.2. Disgust | .29 | .18 | .16\*\*[.05, .27] | .14\*[.02, .25] | .04[-.08, .16] | -.00[-.12, .11] | .07[-.04, .19] | .09 [-.03, .20] |  |  |  |  |  |  |  |  |
| 2.3 Fear | .65 | .20 | **.27**[.16, .38] | .19\*\*[.07, .30] | **.25**[.14, .36] | **.29**[.18, .39] | .19\*\*[.07, .30] | **.43**[.33, .52] | .03[-.09, .15] |  |  |  |  |  |  |  |
| 2.4 Happiness | .47 | .18 | .08[-.04, .19] | .11[-.01, .22] | .18\*\*[.07, .29] | .15\*[.03, .26] | .15\*[.03, .26] | .17\*\* [.05, .28] | .11 [-.01, .22] | **.23**[.12, .34] |  |  |  |  |  |  |
| 2.5 Sadness | .63 | .21 | **.33**[.23, .43] | .20\*\*[.09, .31] | **.25**[.13, .35] | **.27**[.16, .38] | **.29**[.18, .39] | **.50** [.41, .58] | .12 [-.02, .21] | **.36** [.25, .46] | .19\*\* [.07, .30] |  |  |  |  |  |
| Voices |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 3.1 Anger | .73 | .22 | **.32** [.21, .42] | .03[-.09, .14] | .18\*\*[.06, .29] | **.23**[.11, .33] | **.30**[.19, .40] | **.34**[.23, .44] | .15\*\*[.04, .27] | **.28**[.17, .39] | .16\*\*[.04, .27] | **.23**[.12, .34] |  |  |  |  |
| 3.2 Disgust | .77 | .18 | .14\* [.02, .25] | .12\*[.00, .23] | .18\*\*[.07, .29] | .14\*[.02, .25] | .15\*[.04, .26] | .19\*\*[.08, .30] | .11[-.01, .22] | **.22**[.10, .33] | .08[-.04, .20] | .19\*\*[.07, .30] | .18\*\* [.06, .29]  |  |  |  |
| 3.3 Fear | .78 | .18 | .15\* [.04, .27] | .08[-.04, .19] | **.23**[.11, .33] | .12\*[.01, .24] | **.26**[.14, .36] | **.29**[.18, .39] | .12\*[.01, .23] | **.23**[.11, .33] | .17\*\*[.05, .28] | **.21**[.10, .32] | .18\*\* [.07, .29] | **.22** [.11, .33] |  |  |
| 3.4 Happiness | .82 | .22 | .08 [-.04, .20] | .03[-.08, .15] | .12\*[.00, .23] | .16\*\*[.05, .28] | .17\*\*[.05, .28] | .12\*[.00, .23] | -.02[-.13, .10] | .17\*\*[.05, .28] | .16\*\*[.02, .26] | .11[-.01, .22] | **.22** [.11, .33] | -.01 [-.13, .11] | .18\*\* [.06, .29] |  |
| 3.5 Sadness | .86 | .15 | .15\*[.03, .26] | .02[-.09, .14] | .04[-.08, .16] | .15\*[.04, .26] | .13\*[.01, .24] | .19\*\*[.08, .30] | .01[-.11, .12] | .12\*[.00, .23] | -.05[-.16, .07] | **.21**[.09, .32] | .16\*\* [.05, .27] | .09 [-.03, .21] | **.21** [.10, .32] | **.24** [.13, .35] |

Table 1. Descriptive statistics and zero-order correlations between emotions and across expressive domain in Study 1. Skew ranged from -1.47 to 0.55; kurtosis ranged from -0.80 to 1.81.

Note. \* *p* ≤ .05; \*\* *p* < .01; Bold indicates *p* < .001; CI95% are presented in square brackets.

Table 2. Model output for confirmatory factor analyses in Study 1. Note that model fit values are unavailable for non-identified models, and therefore these are denoted by dashes.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Model | Description | χ2 (df) | RMSEA | CFI | AIC |
| 1 | Uncorrelated factors | 304.47 (90) | .09 | .61 | 17344.90 |
| 2 | Single factor | 125.31 (90) | .04 | .94 | 17165.73 |
| 3 | Higher order supramodal | 109.52 (87) | .03 | .96 | 17155.94 |
| 4 | Supramodal visual modalities  | 199.94 (88) | .07 | .80 | 17244.37 |
| 5 | Bifactor | - | - | - | - |

Table 3. Associations between the emotion recognition modality blocks, and measures of autism-like traits (AQ-28), alexithymia (TAS-20) and social anxiety (Mini-SPIN) in Study 1.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Voice | Face | Body | AQ-28 | TAS-20 |
| r | p | r | p | r | p | r | p | r | p |
| Face | **.41** | <.001 |  |  |  |  |  |  |  |  |
| Body | .**46** | <.001 | **.53** | <.001 |  |  |  |  |  |  |
| AQ-28 | .01 | .840 | .01 | .887 | .06 | .337 |  |  |  |  |
| TAS-20 | -.11 | .056 | **-.24** | <.001 | **-.22** | <.001 | **.49** | <.001 |  |  |
| Mini-SPIN | .12 | .039 | -.04 | .559 | .02 | .798 | **.56** | <.001 | **.45** | <.001 |

Note. Bold indicates *p*< .001.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Mean | SD | 1.1 | 1.2 | 1.3 | 1.4 | 1.5 | 2.1 | 2.2 | 2.3 | 2.4 | 2.5 | 3.1 | 3.2 | 3.3 | 3.4 |
| Faces |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 1.1 Anger | .69 | .28 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 1.2 Disgust | .63 | .27 | -.01 [-.15, .12] |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 1.3 Fear | .69 | .29 | .09[-.04, .22] | -.02 [-.16, .11] |  |  |  |  |  |  |  |  |  |  |  |  |
| 1.4 Happiness | .74 | .29 | -.04[-.17, .09] | .02 [-.12, .15] | .01 [-.12, .14] |  |  |  |  |  |  |  |  |  |  |  |
| 1.5 Sadness | .61 | .29 | .19**\*\*** [.06, .32] | .04 [-.09, .18] | .01 [.15, .37] | -.03[-.13, .14] |  |  |  |  |  |  |  |  |  |  |
| Bodies |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 2.1 Anger | .80 | .26 | .10[-.03, .23] | .04[-.09, .18] | .14\*[.01, .27] | .13\*[.00, .26] | .11[-.03, .24] |  |  |  |  |  |  |  |  |  |
| 2.2. Disgust | .39 | .26 | .04[-.09, .17] | .15\*[.01, .27] | .01[-.12, .15] | .10[-.04, .23] | .06[-.08, .19] | .19\*\* [.06, .32] |  |  |  |  |  |  |  |  |
| 2.3 Fear | .74 | .27 | .06[-.08, .19] | .00[-.13, .14] | .11[-.02, .24] | -.01[-.14, .13] | -.04[-.17, .10] | .18\*\*[.04, .30] | -.06[-.19, .08] |  |  |  |  |  |  |  |
| 2.4 Happiness | .67 | .26 | .08[-.05, .21] | -.01[-.14, .13] | .11[-.03, .24] | .07[-.07, .20] | .04[-.09, .18] | .19\*\* [.05, .31] | -.03[-.16, .11] | **.28**[.13, .36] |  |  |  |  |  |  |
| 2.5 Sadness | .77 | .25 | **.25**[.13, .37] | -.03[-.16, .11] | **.24**[.11, .36] | .00[-.13, .14] | .10[-.04, .23] | .11 [-.02, .24] | .13 [-.00, .26] | .08 [-.05, .21] | .02 [-.11, .15] |  |  |  |  |  |
| Voices |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 3.1 Anger | .68 | .28 | .08 [-.05, .21] | -.03[-.16, .11] | -.10[-.04, .23] | .07[-.07, .20] | .03[-.11, .16] | .07[-.06, .20] | .01[-.12, .15] | -.02[-.16, .11] | -.08[-.21, .05] | .19\*\*[.06, .32] |  |  |  |  |
| 3.2 Disgust | .74 | .24 | -.02 [-.15, .11] | .02[-.12, .15] | -.04[-.17, .09] | .13[-.00, .26] | .03[-.11, .16] | .04[-.10, .17] | .09[-.05, .22] | -.01[-.15, .12] | -.07[-.20, .07] | .03[-.11, .16] | .10 [-.03, .23]  |  |  |  |
| 3.3 Fear | .82 | .24 | .15\* [.02, .28] | -.04[-.17, .09] | .21\*\*[.08, .34] | .06[-.07, .19] | .14\*[.00, .26] | **.24**[.11, .36] | .06[-.07, .19] | .06[-.08, .19] | .15\*[.02, .28] | .07[-.07, .20] | -.10 [-.23, .04] | -.08 [-.21, .06] |  |  |
| 3.4 Happiness | .61 | .35 | .06 [-.08, .19] | -.06[-.19, .07] | .01[-.13, .14] | .22\*\*[.09, .35] | -.14\*[-.27, -.00] | -.02[-.15, .11] | .09[-.05, .22] | -.04[-.17, .09] | .06[-.07, .19] | -.02[-.15, .12] | -.05 [-.18, .09] | -.03 [-.16, .10] | -.03 [-.16, .11] |  |
| 3.5 Sadness | .82 | .26 | .03[-.10, .16] | .11[-.03, .24] | .08[-.05, .21] | -.07[-.20, .06] | -.10[-.23, .03] | -.01[-.14, .13] | -.05[-.18, .08] | -.04[-.17, .10] | .10[-.04, .23] | .03[-.10, .16] | .03[-.10, .17] | .17\* [.04, .30] | .03 [-.11, .16] | .16\* [.03, .29] |

Table 4. Descriptive statistics and zero-order correlations between emotions and across expressive domain in Study 2. Skew ranged from -1.44 to 0.21; kurtosis ranged from -1.09 to 1.52.

Note. \* *p* ≤ .05; \*\* *p* < .01; Bold indicates *p* < .001; CI95% are presented in square brackets.

Table 5. Associations between the emotion recognition modality blocks, the Reading the Mind in Films Task (RMFT), the face recognition tasks (GFMT: Glasgow Face Matching Test, CFMT: Cambridge Face Memory Test, and the Mooney Face Test) and the Vocabulary test in Study 2.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Voice | Face | Body | RMFT | GFMT | CFMT | Mooney |
| r | p | r | p | r | p | r | p | r | p | r | p | r | p |
| Face | **.14** | .045 |  |  |  |  |  |  |  |  |  |  |  |  |
| Body | .13 | .054 | **.29** | <.001 |  |  |  |  |  |  |  |  |  |  |
| RMFT | **.17** | .010 | **.26** | <.001 | **.26** | <.001 |  |  |  |  |  |  |  |  |
| GFMT | **.16** | .015 | **.19** | .004 | .12 | .071 | .09 | .176 |  |  |  |  |  |  |
| CFMT | .10 | .152 | **.23** | .001 | .09 | .201 | .13 | .052 | **.28** | <.001 |  |  |  |  |
| Mooney | -.01 | .885 | .06 | .349 | .08 | .247 | **.17** | .012 | **.15** | .023 | .08 | .269 |  |  |
| Vocabulary | **.14** | .046 | **.22** | .001 | **.18** | .009 | **.42** | <.001 | .01 | .897 | **.18** | .007 | **.14** | .041 |

Note. Bold indicates *p* < .05.

Table 6. Model output for confirmatory factor analyses in Study 2. Note that model fit values are unavailable for non-identified models, and therefore these are denoted by dashes.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Model | Description | χ2 (df) | RMSEA | CFI | AIC |
| 1 | Uncorrelated factors | 29.68 (14) | .07 | .82 | 7233.01 |
| 2 | Single factor | 20.70 (14) | .05 | .92 | 7224.02 |
| 3 | Higher order socio-cognitive | 11.88 (13) | .00 | 1.00 | 7217.21 |
| 4 | Bifactor | - | - | - | - |
| 4a | Reduced bifactor with emotion factor | 9.45 (10) | .00 | 1.00 | 7220.77 |
| 4b | Reduced bifactor with face factor | 10.67 (11) | .00 | 1.00 | 7220.00 |
| 5 | Higher order: verbal IQ to socio-cognitive | 36.00 (19) | .06 | .87 | 8062.05 |
| 6 | Higher order: verbal IQ to socio-cognitive and RMFT | 20.21 (18) | .02 | .98 | 8048.26 |
| 7 | Higher order: verbal IQ to socio-cognitive, RMFT and emotion factor | 18.27 (17) | .02 | .99 | 8048.32 |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Mean | SD | 1.1 | 1.2 | 1.3 | 1.4 | 1.5 | 2.1 | 2.2 | 2.3 | 2.4 | 2.5 | 3.1 | 3.2 | 3.3 | 3.4 |
| Faces |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 1.1 Anger | .66 | .28 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 1.2 Disgust | .59 | .32 | .14\* [.03, .24] |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 1.3 Fear | .67 | .29 | .03 [-.09, .16] | .19\*\* [.07, .31] |  |  |  |  |  |  |  |  |  |  |  |  |
| 1.4 Happiness | .70 | .31 | .11 [-.02, .23] | .15\* [.03, .27] | .08 [-.05, .20] |  |  |  |  |  |  |  |  |  |  |  |
| 1.5 Sadness | .59 | .29 | .19**\*\*** [.07, .31] | .02 [-.10, .15] | .09 [-.03, .21] | .14\*[.01, .26] |  |  |  |  |  |  |  |  |  |  |
| Bodies |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 2.1 Anger | .71 | .31 | **.31**[.20, .42] | **.27**[.15, .38] | .20**\*\***[.08, .32] | .10[-.03, .22] | .17\*\*[.05, .29] |  |  |  |  |  |  |  |  |  |
| 2.2. Disgust | .37 | .27 | .08[-.05, .20] | .09[-.03, .21] | .04[-.09, .16] | .04[-.09, .16] | .08[-.05, .20] | **.27** [.15, .38] |  |  |  |  |  |  |  |  |
| 2.3 Fear | .73 | .28 | .14**\***[.02, .26] | .18\*\*[.06, .30] | .18\*\*[.06, .30] | .14\*[.01, .26] | .21\*\*[.09, .33] | **.34**[.23, .45] | .20\*\*[.07, .31] |  |  |  |  |  |  |  |
| 2.4 Happiness | .65 | .29 | .20**\*\***[.07, .31] | .02[-.10, .15] | .09[-.04, .21] | .12[-.00, .24] | .09[-.04, .21] | **.27** [.15, .38] | **.23** [.11, .34] | **.30**[.18, .41] |  |  |  |  |  |  |
| 2.5 Sadness | .67 | .28 | **.23**[.11, .34] | .12\*[.00, .25] | .13\*[.01, .25] | .13\*[.00, .25] | **.25**[.13, .36] | **.37** [.26, .47] | .16\* [.03, .27] | **.34** [.23, .45] | .19\*\* [.06, .30] |  |  |  |  |  |
| Voices |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 3.1 Anger | .32 | .16 | .08 [-.05, .20] | .06[-.06, .18] | .10[-.03, .22] | .05[-.08, .17] | .12[-.01, .24] | .08[-.04, .21] | .07[-.06, .19] | .06[-.06, .19] | .02[-.10, .15] | .12[-.01, .24] |  |  |  |  |
| 3.2 Disgust | .70 | .27 | .05 [-.07, .18] | .10[-.03, .22] | .06[-.06, .19] | .07[-.06, .19] | .07[-.06, .19] | .18\*\*[.06, .30] | .15\*[.02, .27] | .14\*[.01, .26] | .14\*[.01, .26] | .07[-.06, .19] | .02 [-.11, .14]  |  |  |  |
| 3.3 Fear | .79 | .27 | .02 [-.11, .14] | .14\*[.01, .26] | .18\*\*[.06, .30] | .04[-.09, .16] | .02[-.11, .14] | **.31**[.20, .42] | .13\*[.01, .25] | .19\*\*[.07, .31] | .09[-.03, .22] | **.26**[.14, .37] | -.13\* [-.25, -.01] | .16\* [.03, .28] |  |  |
| 3.4 Happiness | .61 | .35 | .07 [-.06, .19] | .19\*\*[.07, .31] | .08[-.04, .21] | .10[.03, .22] | .04[-.08, .17] | .19\*\*[.07, .31] | .08[-.04, .20] | .17\*\*[.05, .29] | .04[-.08, .17] | .14\*[.01, .26] | .03 [-.10, .15] | -.02 [-.14, .11] | .20\*\* [.01, .24] |  |
| 3.5 Sadness | .80 | .27 | .08[-.04, .21] | .20\*\*[-.08, .31] | .10[-.03, .22] | .04[-.08, .17] | .00[-.12, .12] | .19\*\*[.07, .31] | .00[-.12, .12] | .14\*[.01, .26] | .07[-.06, .19] | .14\*[.02, .26] | .10 [-.03, .22] | .10 [-.02, .22] | .11 [-.01, .23] | **.32** [.20, .42] |

Table 7. Descriptive statistics and zero-order correlations between emotions and across expressive domain in Study 3. Skew ranged from -1.12 to 0.22; kurtosis ranged from -1.10 to 2.68.

Note. \* *p* ≤ .05; \*\* *p* < .01; Bold indicates *p* < .001; CI95% are presented in square brackets.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Voice | Face | Body | GFMT | CFMT |
| r | p | r | p | r | p | r | p | r | p |
| Face | **.29** | <.001 |  |  |  |  |  |  |  |  |
| Body | **.38** | <.001 | **.41** | <.001 |  |  |  |  |  |  |
| GFMT | -.01 | .902 | **.22** | .004 | **.28** | <.001 |  |  |  |  |
| CFMT | .02 | .706 | **.30** | <.001 | **.19** | .002 | **.37** | <.001 |  |  |
| General Intelligence | **.38** | <.001 | **.32** | <.001 | **.46** | <.001 | .08 | .202 | .10 | .106 |

Table 8. Associations between the emotion recognition modality blocks, the face recognition tasks (GFMT: Glasgow Face Matching Test, and CFMT: Cambridge Face Memory Test, and the Alice Heim general intelligence test in Study 3.

Note. Bold indicates p < .05.

Table 9. Model output for confirmatory factor analyses in Study 3. Note that model fit values are unavailable for non-identified models, and therefore these are denoted by dashes.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Model | Description | χ2 (df) | RMSEA | CFI | AIC |
| 1 | Uncorrelated factors | - | - | - | - |
| 2 | Single factor | 163.93 (14) | .21 | .63 | 7753.85 |
| 3 | Higher order socio-cognitive | 422.92 (21) | .06 | .97 | 7617.56 |
| 4 | Bifactor | - | - | - | - |
| 4a | Reduced bifactor with emotion factor | 6.89 (9) | .00 | 1.00 | 7606.81 |
| 4b | Reduced bifactor with face factor | - | - | - | - |
| 5 | Higher order: intelligence to socio-cognitive | 60.66 (17) | .10 | .91 | 9571.33 |
| 6 | Higher order: intelligence to emotion factor | 28.81 (17) | .05 | .98 | 9539.48 |
| 7 | Higher order: intelligence to face factor | 103.60 (17) | .14 | .82 | 9614.27 |
| 8 | Higher order: intelligence to socio-cognitive and emotion factor | 26.97 (16) | .05 | .98 | 9539.64 |
| 9 | Reduced bifactor: intelligence to emotion | 10.49 (14) | .00 | 1.00 | 9527.17 |

Figure 1. Schematic representation of the five tested models in Study 1.



Figure 2. Graphical representation of the final model of Study 1 parameter estimates (and 95% confidence intervals). All path coefficients in bold were significant at *p*<.02.

Face Anger

Face Disgust

Face Fear

Face Happiness

Face Sadness

Body Anger

Body Disgust

Body Fear

Body Happiness

Body Sadness

Voice Anger

Voice Disgust

Voice Fear

Voice Happiness

Voice Sadness

**.92**

[.80, 1.0]

**.91**

[.80, 1.0]

**.87**

[.74, 1.0]

**.59**

[.49, .69]

**.46**

[.34, .58]

**.30**

[.17, .42]

**.46**

[.35, .58]

**.48**

[.37, .59]

**.54**

[.43, .65]

**.73**

[.65, .81]

.15

[.02, .28]

**.59**

[.50, .68]

**.30**

[.18, .42]

**.65**

[.56, .74]

**.54**

[.42, .66]

**.35**

[.22, .48]

**.32**

[.19, .45]

**.34**

[.21, .47]

Figure 3. Graphical representation of the final model of Study 2 parameter estimates (and 95% confidence intervals). All path coefficients in bold were significant at *p*<.02. Note RMFT = Reading the Mind in Films Task; GFMT = Glasgow Face Matching Test; CFMT = Cambridge Face Memory Test.

**.55**

[.35, .75]

Voice

Face

Body

RMFT

GFMT

Mooney

CFMT

**.48**

[.32, .64]

**.58**

[.42, .74]

**.49**

[.33, .65]

**.50**

[.31, .70]

**.24**

[.06, .42]

**.76**

[.60, .92]

**.76**

[.60, .92]

**.30**

[.13, .46]

Figure 4. Graphical representation of the final model of Study 2 parameter estimates (and 95% confidence intervals) and inclusion of verbal intelligence. All path coefficients in bold were significant at *p*<.02. Note RMFT = Reading the Mind in Films Task; GFMT = Glasgow Face Matching Test; CFMT = Cambridge Face Memory Test.

**.48**

[.29, .66]

Verbal Intelligence

Voice

Face

Body

RMFT

GFMT

Mooney

CFMT

**.36**

[.19, .58]

**.60**

[.44, .77]

**.48**

[.32, .64]

**.54**

[.34, .73]

**.25**

[.07, .43]

**.30**

[.17, .43]

**.78**

[.61, .94]

**.78**

[.61, .94]

**.40**

[.22, .59]

**.29**

[.12, .46]

Figure 5. Graphical representation of the higher-order model of Study 3 parameter estimates (and 95% confidence intervals). All path coefficients in bold were significant at *p*<.001. Note GFMT = Glasgow Face Matching Test; CFMT = Cambridge Face Memory Test.

Voice

Face

Body

GFMT 1

CFMT 1

**.46**

[.33, .59]

GFMT 2

CFMT 2

**.59**

[.46, .72]

**.74**

[.61, .88]

**.58**

[.37, .78]

**.71**

[.50, .92]

**.44**

[.27, .61]

**.50**

[.33, .67]

**.70**

[.57, .84]

**.70**

[.57, .84]

.22

[-.11, .55]

**.60**

[.50 .70]

Figure 6. Graphical representation of the final higher-order model of Study 3 parameter estimates (and 95% confidence intervals) and the inclusion of general intelligence. All path coefficients in bold were significant at *p*<.001. Note RMFT = Reading the Mind in Films Task; GFMT = Glasgow Face Matching Test; CFMT = Cambridge Face Memory Test.

Voice

Face

Body

GFMT 1

CFMT 1

**.50**

[.38, .61]

GFMT 2

CFMT 2

General intelligence

**.57**

[.46, .68]

**.74**

[.64, .84]

**.59**

[.37, .81]

**.75**

[.52, .98]

**.42**

[.26, .59]

**.48**

[.31, .65]

**.55**

[.40, .69]

**.57**

[.44, .69]

.15

[-.06, .36]

**.73**

[.58, .87]

.18

[-.23, .59]

**.61**

[.51, .70]

**Supplementary Materials**

Figure S1. Graphical representation of the reduced bifactor model of Study 3 parameter estimates (and 95% confidence intervals) and the inclusion of general intelligence. All path coefficients in bold were significant at *p*<.001.

CFMT 2

Voice

Face

Body

GFMT 1

CFMT 1

GFMT 2

**.63**

[.45, .81]

**.44**

[.29, .59]

**.58**

[.42, .75]

.24

[-.00, .49]

**.59**

[.49, .69]

.03

[-.13, .19]

**.41**

[.26, .56]

**.39**

[.24, .54]

**.59**

[.42, .76]

**.67**

[.50, .84]

**.46**

[.31, .61]

**.52**

[.37, .66]

Figure S2. Graphical representation of the final reduced bifactor model of Study 3 parameter estimates (and 95% confidence intervals) and the inclusion of general intelligence. All path coefficients in bold were significant at *p*<.001.

General intelligence

CFMT 2

Voice

Face

Body

GFMT 1

CFMT 1

GFMT 2

**.66**

[.55, .77]

.12

[-.03, .28]

**.59**

[.47, .72]

**.43**

[.30, .56]

**.63**

[.51, .75]

.22

[-.05, .49]

**.60**

[.50, .69]

-.02

[-.16, .13]

**.37**

[.22, .51]

**.33**

[.19, .46]

**.60**

[.42, .77]

**.69**

[.52, .86]

**.45**

[.30, .60]

**.51**

[.36, .65]