

## The inside out model of emotion recognition

Keating, Connor Tom; Cook, Jennifer

DOI:

[10.31234/osf.io/476dk](https://doi.org/10.31234/osf.io/476dk)

[10.1038/s41598-023-48469-8](https://doi.org/10.1038/s41598-023-48469-8)

License:

Creative Commons: Attribution (CC BY)

*Document Version*

Publisher's PDF, also known as Version of record

*Citation for published version (Harvard):*

Keating, CT & Cook, J 2023, 'The inside out model of emotion recognition: how the shape of one's internal emotional landscape influences the recognition of others' emotions', *Scientific Reports*.

<https://doi.org/10.31234/osf.io/476dk>, <https://doi.org/10.1038/s41598-023-48469-8>

[Link to publication on Research at Birmingham portal](#)

### General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

### Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact [UBIRA@lists.bham.ac.uk](mailto:UBIRA@lists.bham.ac.uk) providing details and we will remove access to the work immediately and investigate.



## OPEN The inside out model of emotion recognition: how the shape of one's internal emotional landscape influences the recognition of others' emotions

Connor Tom Keating & Jennifer Louise Cook

Some people are exceptional at reading emotional expressions, while others struggle. Here we ask whether the way we experience emotion “on the inside” influences the way we expect emotions to be expressed in the “outside world” and subsequently our ability to read others’ emotional expressions. Across multiple experiments, incorporating discovery and replication samples, we develop EmoMap (N = 20; N = 271) and ExpressionMap (N = 98; replication N = 193) to map adults’ experiences of emotions and visual representations of others’ emotions. Some individuals have modular maps, wherein emotional experiences and visual representations are consistent and distinct—anger looks and feels different from happiness, which looks and feels different from sadness. In contrast, others have experiences and representations that are variable and overlapping—anger, happiness, and sadness look and feel similar and are easily confused for one another. Here we illustrate an association between these maps: those with consistent and distinct *experiences* of emotion also have consistent and distinct *visual representations* of emotion. Finally (N = 193), we construct the *Inside Out Model of Emotion Recognition*, which explains 60.8% of the variance in emotion recognition and illuminates multiple pathways to emotion recognition difficulties. These findings have important implications for understanding the emotion recognition difficulties documented in numerous clinical populations.

Some people are exceptional at navigating the social world: the considerate concierge rapidly reads facial expressions and anticipates every desire; the perceptive companion accurately detects the sadness behind their friend’s smile; the skilled negotiator notices a telling tightness around the eyes and knows just the right time to apply pressure. Other individuals struggle: as Parkinson’s Disease progresses, people with this condition increasingly report challenges with reading others’ emotional expressions<sup>1</sup>, and similar difficulties predict negative social and wider health outcomes across a range of psychiatric and mental health conditions<sup>2–4</sup>. Despite clear individual differences in the ability to read others’ emotional expressions, little is known about why these individual differences exist. Here we ask whether individual differences in navigating the social world of others’ facial expressions are related to individual differences in the shape of one’s own *internal* emotional landscape. In other words, is there a relationship between our experience of emotion “on the inside” and our ability to identify emotions in the “outside world”?

Internal “maps” of concepts—such as personality traits—can exert a considerable influence on the judgments we make about others. Stolier et al.<sup>5</sup> for instance, mapped internal conceptual-trait maps by asking participants to rate the similarity of 13 different personality traits. They also mapped representations of how these traits are depicted on people’s faces by asking participants to rate various face images with respect to these 13 traits. Both internal (semantic) conceptual maps and external maps of facial representations, tended to exhibit a modular structure with particular traits—such as aggressive, mean, dominant and egotistical—clustering together. Importantly, the shape of an individual’s map of facial representations was highly correlated with the shape of their internal conceptual landscape, such that a perceiver who believed aggression and dominance to be closely related in conceptual space would be more likely (compared to a perceiver with a weak link between the two concepts) to see an aggressive face as dominant. Thus, Stolier and colleagues illustrate that, for trait judgements, internal conceptual maps and judgements we make about others in the outside world are tightly related.

University of Birmingham, Birmingham, UK. ✉ email: cxx655@bham.ac.uk

Stolier and colleagues' work pertains to traits. Here we focus on emotions. Preliminary evidence provides initial support for a link between the experience and recognition of emotion. Israelashvili and colleagues<sup>6</sup> for example, illustrated that individuals who are good at differentiating their own experiences of distinct emotions are more accurate in reading others' emotional facial expressions. Nevertheless, although preliminary evidence indicates that individuals who are better able to identify how they feel "on the inside" are also better able to recognize emotions in the "outside world", it is unclear why this relationship exists. Afterall, recognizing *one's own* emotions primarily depends upon the labelling of *internal signals*, whereas recognizing *others'* emotions principally consists of categorizing *incoming sensory information*. The psychological mechanisms supporting superior emotion recognition in individuals with superior (own) emotion differentiation are currently unknown.

The face identity literature provides a candidate mechanism: studies from this field have illustrated that individuals who are good at face identity recognition tend to have robust visual representations (also referred to as templates and/or abstracted structural representations) of others' identities, in their minds eye<sup>7-9</sup>. Such representations are thought to be constructed via experience wherein exposure to different views of a face updates the abstracted structural representation of this identity and, over time, the representation comes to emphasize diagnostic aspects of the face (that differentiate this face from another) and minimize non-diagnostic aspects. Signal detection theory (see<sup>10</sup>) also tells us that distinguishing between signal and noise (e.g., correct and incorrect facial identities) is easier if the signal and noise distributions are distinct and precise—when these channels are not overlapping, and when they are *consistent* across numerous instances or samples (i.e., narrow). In line with this, Etchells et al.<sup>9</sup> found that participants were better at recognizing faces from a novel view when they had built up a more precise representation of that face from multiple views, relative to a single view, during a preceding learning phase. Furthermore, it is well documented that faces that are more overlapping in appearance are more difficult to differentiate<sup>11</sup>. Therefore, the face identity literature raises the hypothesis that individuals who are adept at reading others' emotions will have *consistent* and *distinct* visual representations (in their 'mind's eye') of emotional facial expressions. This hypothesis is yet to be tested.

If we are to understand why people who are better at recognizing *others'* emotions tend to be good at identifying *their own*, and if this is related to the consistency of visual representations of *others'* emotional expressions, we must also explain why representations would be more consistent and distinct for individuals who are better able to differentiate their *own* emotions. Models of conceptual learning suggest that robust concepts facilitate learning: having a (semantic) concept that a table has a flat top and four legs encourages a learner to focus on these invariant features when encountering new table exemplars and ignore variant features such as color or texture<sup>12,13</sup>, thus minimizing within-category differences and maximizing between-category differences<sup>14</sup>. Similarly, having consistent and distinct concepts of *one's own* emotions (which may be multidimensional including semantic, interoceptive and sensory information<sup>15-20</sup>) may encourage a learner to focus on invariant features of facial expressions and ignore between expression variation, thus encouraging the formation of consistent and distinct visual representations of *others'* facial expressions. However, despite theoretical justification for a link between the experience and representation of emotion, research has not yet tested this idea.

Here we ask whether the experience of emotion "on the inside" influences the way in which one represents the dynamic emotional facial expressions that one would encounter in the "outside world" and whether this, in turn, affects emotion recognition accuracy. Specifically, we predict that some individuals will have internal emotion maps with a clear modular structure, wherein emotional experiences are *consistent* and *distinct*: happiness feels very different from anger, which feels very different from sadness. We predict that these individuals will also have *consistent* and *distinct* representations of the way in which emotions are expressed on others' faces and, correspondingly, will be adept at recognizing expressions. Other individuals, however, may have *variable* and *overlapping* experiences of emotion wherein anger, happiness and sadness feel relatively similar and are easily confused for one another. We predict that these individuals will have more *variable* and *overlapping* visual representations of others' expressions such that, in their mind's eye, anger, happiness and sadness look relatively similar. Thus, resulting in emotion recognition difficulties.

Across a series of experiments, we first develop and validate "EmoMap", a novel method to map the shape of individuals' emotional experience landscapes (Expt 1). Second, we develop "ExpressionMap" to map the landscape of participants' visual representations of emotional expressions (Expt 2). Following this, we test for a mapping between the experience of emotion "on the inside" and representations of the way emotions are expressed in the "outside world". That is, we ask whether those with modular internal emotion maps, who have consistent and distinct experiences of anger, happiness and sadness, also tend to have consistent and distinct visual representations of angry, happy and sad facial expressions (note that these emotions were selected as they correspond to different regions in the circumplex model of emotion<sup>21</sup>, varying in both and valence). Throughout these analyses, we control for clinically relevant demographic factors known to be associated with the experience and perception of emotion (e.g., the level of autistic traits, the level of alexithymic traits, and non-verbal reasoning ability<sup>22-28</sup>) to ensure that any relationships we discover exist even after accounting for these variables. Finally, we assess the contribution of the consistency and differentiation (i.e., distinctness) of emotional experiences and representations to the recognition of anger, happiness and sadness, and use structural equation modelling to construct the 'Inside Out Model' of emotion recognition; a model which provides insight into the psychological mechanisms by which one's experience of emotions "on the inside" influences one's ability to identify emotions in the "outside world".

## Results

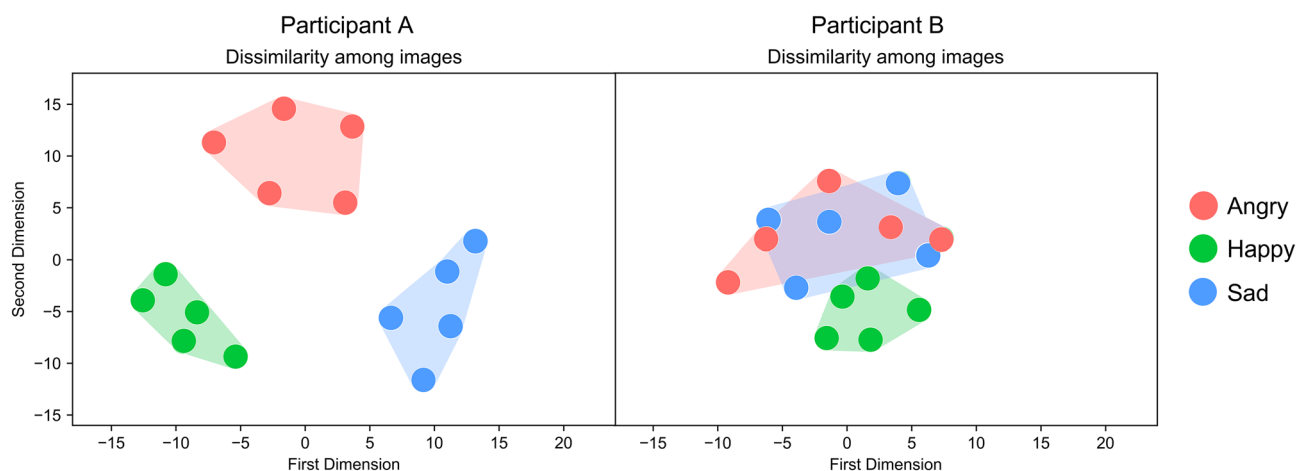
### Study 1: developing EmoMap

Participants ( $N=271$ ) completed our EmoMap paradigm—a two-part task that assesses the differentiation and consistency of emotional experiences. In the first part, on each trial, participants viewed pairs of images (from

the Nencki Affective Picture System<sup>29</sup>) each known to selectively induce either anger, happiness or sadness<sup>30</sup>, and were asked to rate how similar the emotions evoked by the images were on a scale from zero, 'Not at all similar', to ten, 'Very similar' (to 4 decimal places). These similarity scores were then transformed into distance scores via multidimensional scaling, a statistical technique that represents objects (emotional images, lexical items) as points in multidimensional space, wherein close similarity between objects corresponds to small distances between the points in the representation. Distance scores were then used to (a) calculate the mean distances between (e.g., distance between angry and happy clusters, angry and sad clusters, and happy and sad clusters) and within emotion clusters, and (b) plot multidimensional scaling maps.

The multidimensional scaling maps confirmed that the internal emotional landscape had a modular structure for some participants (Fig. 1, left panel) and a less modular, more overlapping, structure for others (Fig. 1, right panel). EmoMap was validated by illustrating that individuals high in alexithymic traits, who by definition have difficulties differentiating their own emotions<sup>31</sup>, tended to have emotional landscape maps with a less modular, more overlapping structure, whereas those low in alexithymic traits had modular emotional landscapes. That is, linear mixed effects models predicting mean distance between clusters and mean distance within clusters with TAS score, AQ score, and non-verbal reasoning ability (clinically relevant demographic variables known to be associated with the experience and perception of emotion; e.g.,<sup>22–28</sup>), and with subject number as a random intercept revealed alexithymic traits as a significant negative predictor of distance *between* emotion clusters [ $F(1,267) = -5.92, p < 0.05$ ] and distance *within* emotion clusters [ $F(1,267) = -6.16, p < 0.05$ ]. In general, greater overlap was seen between anger and sadness [mean distance (SEM) = 14.39 (0.21)], than happiness and anger [mean distance (SEM) = 20.79 (0.29)], and happiness and sadness [mean distance (SEM) = 20.70 (0.29)] in participants' internal emotional landscapes (see Supplementary Information A for a full discussion). These results validate EmoMap by confirming that individuals who, by definition, have difficulties differentiating their own emotions exhibit higher EmoMap emotion confusion as indexed by smaller distances between- and within-emotion clusters (suggesting they have difficulties differentiating distinct and more similar emotional states).

In the second part of EmoMap, on each trial participants were required to make decisions about three images (also from the Nencki Affective Picture System<sup>29</sup>). There were four conditions: one non-emotional control condition, and three emotional experimental conditions exploring the experience of anger, happiness and sadness respectively. Participants completed the control condition first. In this condition, participants were required to select which of the three (emotionally neutral) images they found most colorful using their mouse cursor. Two of these images were in color and one was in grayscale, thus serving as an attention check. Following this, participants completed the three experimental conditions in a random order. In these conditions, participants were required to select which of the three images made them feel most angry, happy or sad using their mouse cursor (i.e., in the 'angry condition' participants would have to decide which image made them most angry). As in the control condition, there was a 'trap' image on each trial such that two of the images were strong inducers of the target emotion (e.g., sadness), and one was a strong inducer of another emotion (e.g., happiness), thus serving as an attention check. Emotional consistency was calculated, for each emotion, based on the logical consistency of decision-making: if a participant selected image A over image B ( $A > B$ ) and image B over image C ( $B > C$ ), but then selected image C over image A ( $C > A$ ), this would be considered an inconsistent decision and would result in a reduction in their consistency score (see Methods for further details on scoring). Consistency requires participants to differentiate precisely between the intensity of emotion evoked by each image. Therefore, inconsistent decisions are likely to stem from inconsistencies in an individual's emotional experience across repeated instances.



**Figure 1.** Examples of consistent and distinct (left), and variable and overlapping (right) emotional landscapes. The dimensions illustrated here may somewhat reflect the two dimensions outlined in the circumplex model of affect<sup>21</sup>—arousal and valence. The first dimension may correspond to valence, with high values reflecting negative valence and low values reflecting positive valence (see left). The second dimension may correspond to arousal; high scores reflect high activation, and low scores reflect low activation (see left). This may be an appropriate interpretation of the internal emotional landscape for Participant A (left).

Using scores from this task, we aimed to determine whether there is a link between the consistency and differentiation of emotional experiences. Our results illustrate that individuals with modular landscapes are more likely to have consistent emotional experiences, whereas those with more overlapping emotional landscapes have less reliable emotional experiences. That is, a linear mixed effects model of emotional consistency as a function of between-cluster distances, within-cluster distances, the interaction between emotion and between-cluster distances, the interaction between emotion and within-cluster distances (independent variables), AQ, TAS, non-verbal reasoning and color (control) consistency (control variables), with subject number as a random intercept revealed that emotional consistency was positively predicted by between-cluster distances [ $F(1,1786.1) = 9.58$ ,  $p < 0.01$ ], and negatively predicted by within-cluster distances [ $F(1,785.9) = -10.30$ ,  $p < 0.01$ ]. Since the emotion that was displayed (angry, happy or sad) did not interact with between- or within-cluster distances to predict emotional consistency, our results suggest that those with larger distances *between* clusters and smaller distances *within* their emotion clusters typically had greater emotional consistency for anger, happiness and sadness. Emotional consistency was also positively predicted by non-verbal reasoning ability [ $F(1,264) = 12.83$ ,  $p < 0.001$ ] but not by any other variables, including color control consistency [all  $p > 0.05$ ]. Hence, our results demonstrate that the distances between and within emotion clusters predict the consistency of emotional experiences.

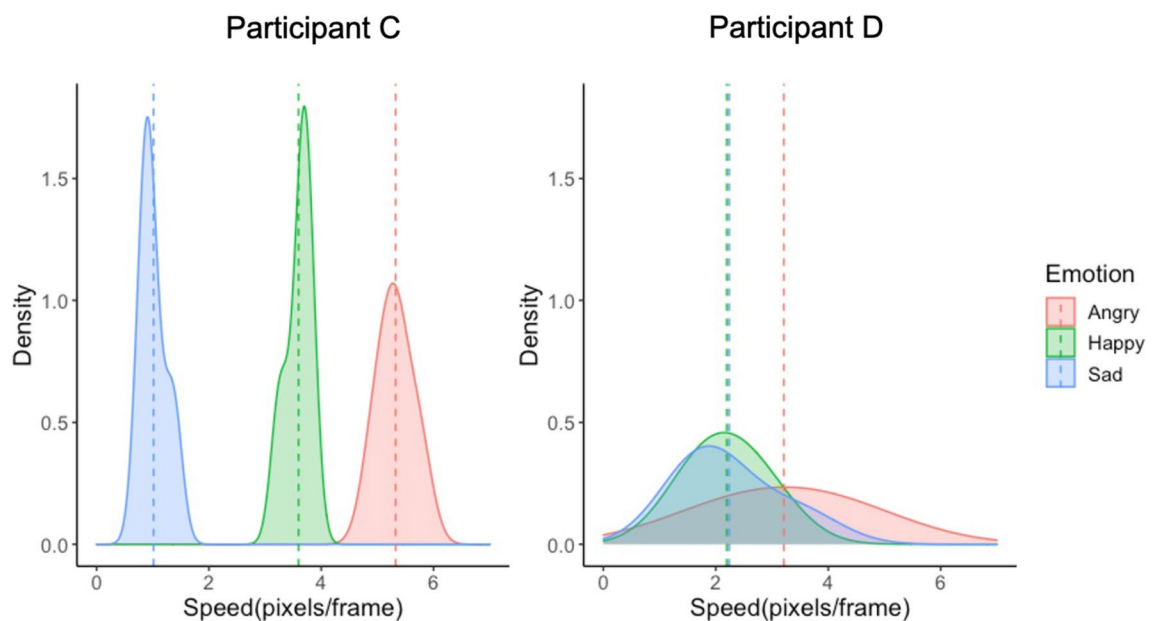
### Study 2: developing ExpressionMap

To map visual representations of the external expression landscape, participants ( $N = 98$ ; replication  $N = 193$ ) completed our ExpressionMap paradigm. On each trial participants were asked to move a dial to change the speed of an emotional point light display of the face (a PLF) until it matched the speed they typically associated with an angry, happy or sad expression. That is, participants were matching the speed of the displayed PLF to their visual representation of that expression. The consistency of visual representations was indexed as the standard deviation of the speeds attributed to each repetition of the angry, happy and sad expressions respectively, multiplied by -1 (see “Methods” for full details). Mean representational consistency was calculated by taking a mean of the consistency scores for the angry, happy and sad PLFs. In addition, this task also provides an index of the ‘distance’ between emotions in participants’ visual representations of facial expressions. Distance scores were calculated as the absolute difference in the speeds attributed to two different emotions. For example, to calculate distance between happy and angry, we subtracted the mean speed attributed to happy PLFs from the mean speed attributed to angry PLFs, and then took the absolute value. Mean distance was calculated by taking a mean of the scores for the angry-happy, angry-sad, and happy-sad distances.

To visualize representations of the external emotional landscape, we produced density plots displaying the speeds attributed to angry, happy and sad expressions respectively. Density plots confirmed that for some individuals, visual representations of emotion are *consistent* and *distinct* (Fig. 2, left panel), and for others they are *variable* and *overlapping* (Fig. 2, right panel). Across participants, the consistency and differentiation of such representations differed as a function of emotion/emotion pair—these results are reported in Supplementary Information A as they are outside the scope of the current study.

*Mapping between the experience of emotion “on the inside” and representations of emotional expressions in the “outside world”*

A subset of participants ( $N = 193$ ) completed both EmoMap and ExpressionMap. To probe the existence of a mapping between the experience of emotion “on the inside” and representations of the way emotions are



**Figure 2.** Examples of consistent and distinct (left), and variable and overlapping (right) visual emotion representations.

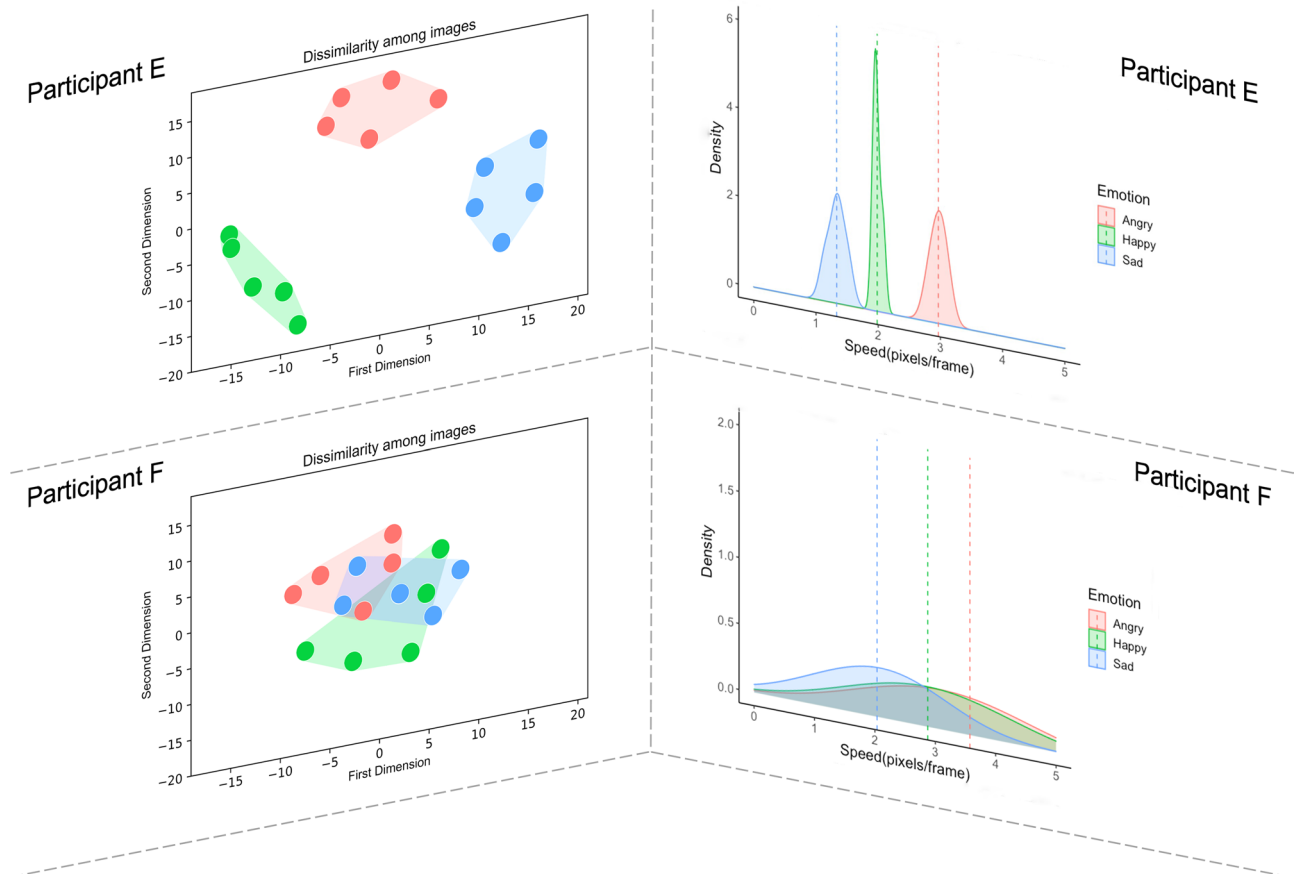


expressed in the “outside world”, we constructed two separate linear mixed effects models to predict metrics of ExpressionMap (representational consistency and distance between representations) from metrics derived from EmoMap (emotional consistency, distance between emotion clusters, respectively), with AQ score, TAS score, non-verbal reasoning, and color (control) consistency as control variables, and with subject number as a random intercept. Representational consistency was positively predicted by emotional consistency [ $F(1,186) = 5.15$ ,  $p < 0.05$ ] but not colour control consistency [ $p > 0.05$ ]: individuals with more consistent experiences of emotion also had more consistent visual representations of emotion (while the consistency of colorfulness judgments did not contribute to the consistency of visual representations). Non-verbal reasoning was also a significant predictor of representational consistency [ $F(1,186) = 30.71$ ,  $p < 0.001$ ]: those with higher non-verbal reasoning had greater representational consistency. In addition, distance between emotion representations was predicted by distance between emotion clusters [ $F(1,186) = 8.19$ ,  $p < 0.01$ ]: those with more distinct experiences of emotion also had more distinct representations. Thus, consistency and differentiation within internal emotional landscapes is linked to consistency and differentiation in visual representations of the external world (even after controlling for relevant participant demographics; see Fig. 3).

#### Predicting emotion recognition ability

The above analyses illustrate a mapping between the experience of emotion “on the inside” and visual representations of the way emotions are expressed in the “outside world”, but how do these inside and outside maps influence emotion recognition accuracy? To answer this question, we first focused on asking how the shape of ExpressionMaps relate to individual differences in emotion recognition as indexed by our previously validated PLF Emotion Recognition Task<sup>22,32</sup>. On each trial in this task, participants viewed an angry, happy or sad PLF and rated the extent to which the expression looked angry, happy and sad. Emotion recognition accuracy was calculated as the correct emotion rating minus the mean of the two incorrect emotion ratings.

Building on the face identity literature<sup>7–9</sup> and principles of signal detection theory<sup>10</sup>, our a priori hypothesis was that emotion recognition accuracy would be positively predicted by the consistency of, and distance between, emotion representations. To test this, we constructed a linear mixed effects model with accuracy as the outcome variable, representational consistency, distance between emotion representations, the interaction between representational consistency and distance, AQ score, TAS score, and non-verbal reasoning as predictors (clinically



**Figure 3.** A diagram demonstrating that consistency and differentiation within internal emotional landscapes (left) is linked to consistency and differentiation in visual models of the external world (right). Figure top shows the modular emotion and representation maps of one participant. Figure bottom shows the overlapping emotion and representation maps of another participant.

relevant participant characteristics known to be involved in the experience and perception of emotion; e.g.,<sup>22–28</sup>), and subject number as a random intercept. Across both our original ( $N=98$ ) and replication ( $N=193$ ) sample, representational consistency was a significant positive predictor of accuracy [original sample:  $F(1,91)=4.19$ ,  $p<0.05$ ; replication sample:  $F(1,186)=13.86$ ,  $p<0.001$ ; see Fig. 4]: those with more consistent visual emotion representations typically achieved higher accuracy (i.e. identified the emotion that the actor intended to convey) on the PLF Emotion Recognition Task. In conflict with our hypothesis, accuracy was not predicted by distance in either sample [all  $p>0.05$ ]. There were also no other significant predictors of accuracy across both samples [all  $p>0.05$ ].

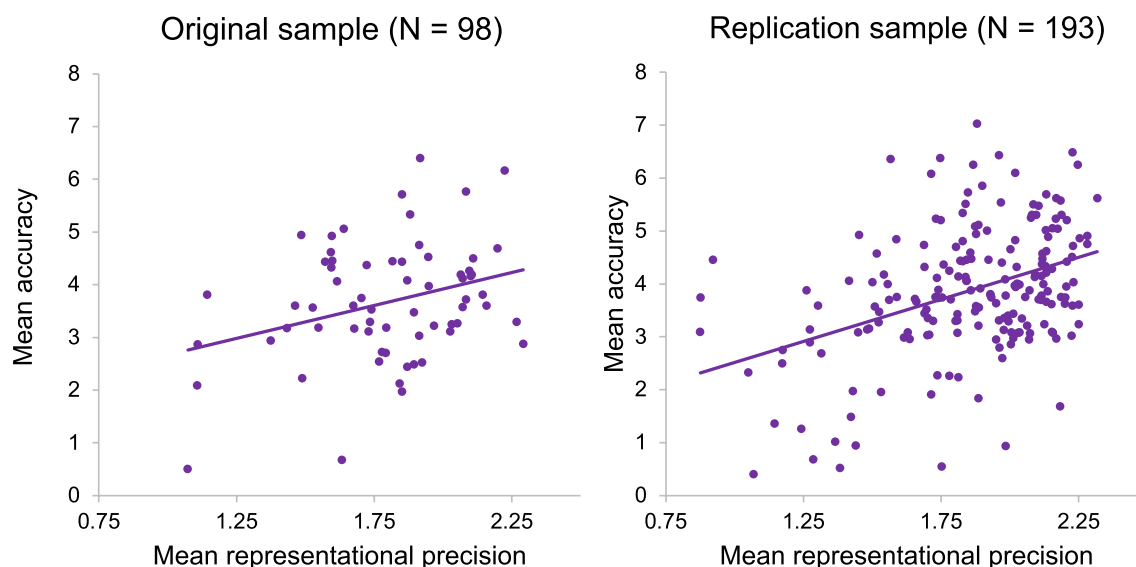
Since it is likely that emotion recognition is contingent not only on the clarity of emotion representations but also on the ability to match a displayed expression to one's visualization, we also included a visual matching task in our battery. This task assesses how well participants can visually match the speed of one expression to another displayed expression. Each trial began with a PLF stimulus video on the left-hand side of the screen. After this video had played once, the same PLF stimulus video also appeared on the right-hand side of screen (moving at a random speed) and continued to play in a loop. Participants were instructed to “move the dial to change the speed of the video on the right until it matches the speed of the video on the left”. Consequently, participants were visually matching the speed of one PLF to another. Deviation scores (the distance between the speeds of the two animations) comprised the absolute value of the percentage speed attributed to the leftward expression minus that attributed to the rightward expression. Mean deviation scores comprised a mean of all of the absolute deviation scores. Higher deviation scores represented greater difficulties matching the two expressions.

Subsequently, visual matching difficulty and the interaction between representational consistency and matching difficulty were added to the linear mixed effects model described above. In our larger sample, we found that the main effect of representational consistency on emotion recognition accuracy was moderated by matching difficulty [ $F(1,184)=12.26$ ,  $p<0.001$ ]. To unpack this interaction, we conducted a median half split analysis, dividing participants into a high matching group (matching deviation scores  $<27.75\%$ ) and a low matching group (matching deviation scores  $>27.75\%$ ). Representational consistency was only a significant predictor of accuracy for those with lower matching ability [ $F(1,89)=7.16$ ,  $p<0.01$ ], and not those with higher matching ability [ $F(1,90)=0.44$ ,  $p=0.507$ ] (see Fig. 5). This interaction was also evident in our original sample [low matching:  $F(1,42)=4.18$ ,  $p<0.05$ ; high matching:  $F(1,42)=0.44$ ,  $p=0.513$ ]. Hence, across both samples, for participants with a lower ability to match expressions, representational consistency was a significant predictor of emotion recognition ability. This potentially indicates that when one's ability to match two representations is compromised, having clear and consistent visual representations becomes particularly important.

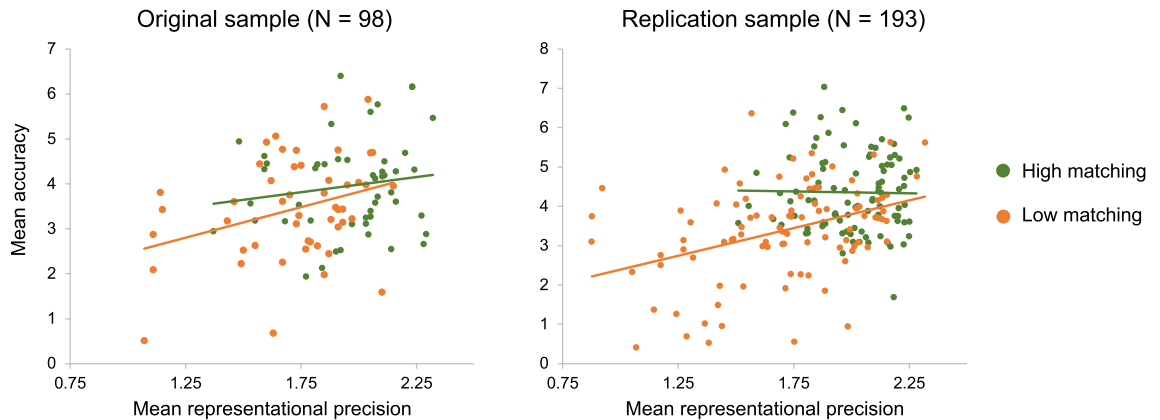
In sum, emotion recognition ability is predicted by the consistency of imagined visual emotion representations and one's matching ability, such that—for individuals with lower ability to match two visually displayed expressions—the more consistent one's representations the better one's emotion recognition accuracy.

#### Building the inside out model of emotion recognition ( $N=193$ )

For the following analyses, we focused on the 193 participants that had completed all four tasks (EmoMap, ExpressionMap, Visual Matching and PLF Emotion Recognition), thus allowing us to build a comprehensive model incorporating the experience, representation and recognition of emotion. Model construction comprised a four-step process. First, since we had many potential variables of interest, we determined their relative importance for emotion recognition using a random forests analysis<sup>33</sup> employing the Boruta wrapper algorithm<sup>34</sup>. In this analysis, representation matching, matching difficulty, representational consistency, distance between



**Figure 4.** The relationship between mean accuracy and mean representational consistency in original sample (left [ $R=0.311$ ,  $p=0.002$ ,  $BF_{10}=15.32$ ]) and replication sample (right [ $R=0.399$ ,  $p<0.001$ ,  $BF_{10}=1.21e^6$ ]).



**Figure 5.** The relationship between mean accuracy and representational consistency within the high (original sample:  $R = 0.148$ ,  $p = 0.310$ ,  $BF_{10} = 0.294$ ; replication sample:  $R = 0.004$ ,  $p = 0.972$ ,  $BF_{10} = 0.13$ ) and low matching groups (original sample:  $R = 0.324$ ,  $p < 0.05$ ,  $BF_{10} = 2.18$ ; replication sample:  $R = 0.379$ ,  $p < 0.001$ ,  $BF_{10} = 176.06$ ).

emotion clusters, and emotional consistency were deemed important for emotion recognition. Here, ‘representation matching’ reflects the interaction between representational consistency and matching difficulty, which was found to be a significant predictor of emotion recognition in our previous analyses. ‘Representation matching’ was computed by multiplying the representational consistency scores for angry, happy and sad expressions with their corresponding matching difficulty scores (e.g., angry representational consistency  $\times$  angry matching difficulty; happy representational consistency  $\times$  happy matching difficulty; sad representational consistency  $\times$  sad matching difficulty). Higher representation matching scores indicate superior representational consistency, matching ability, or both. Following our random forests analysis, we added variables classified as “important” into a structural equation model predicting emotion recognition accuracy, sequentially (starting with the most important variable), until there was no longer a significant improvement (or our goodness of fit index exceeded the specified threshold). Third, to determine the most mathematically plausible path directions in our structural equation model, we systematically reversed each path and compared the Bayesian Information Criterion (BIC) scores for the original and reversed models (see Supplementary Information B for the steps listed above). Lastly, we built one final structural equation model in which we included the path directions that were mathematically most plausible. There was strong to very strong evidence that this final model was better than all previous models (BIC difference  $> 6$ ).

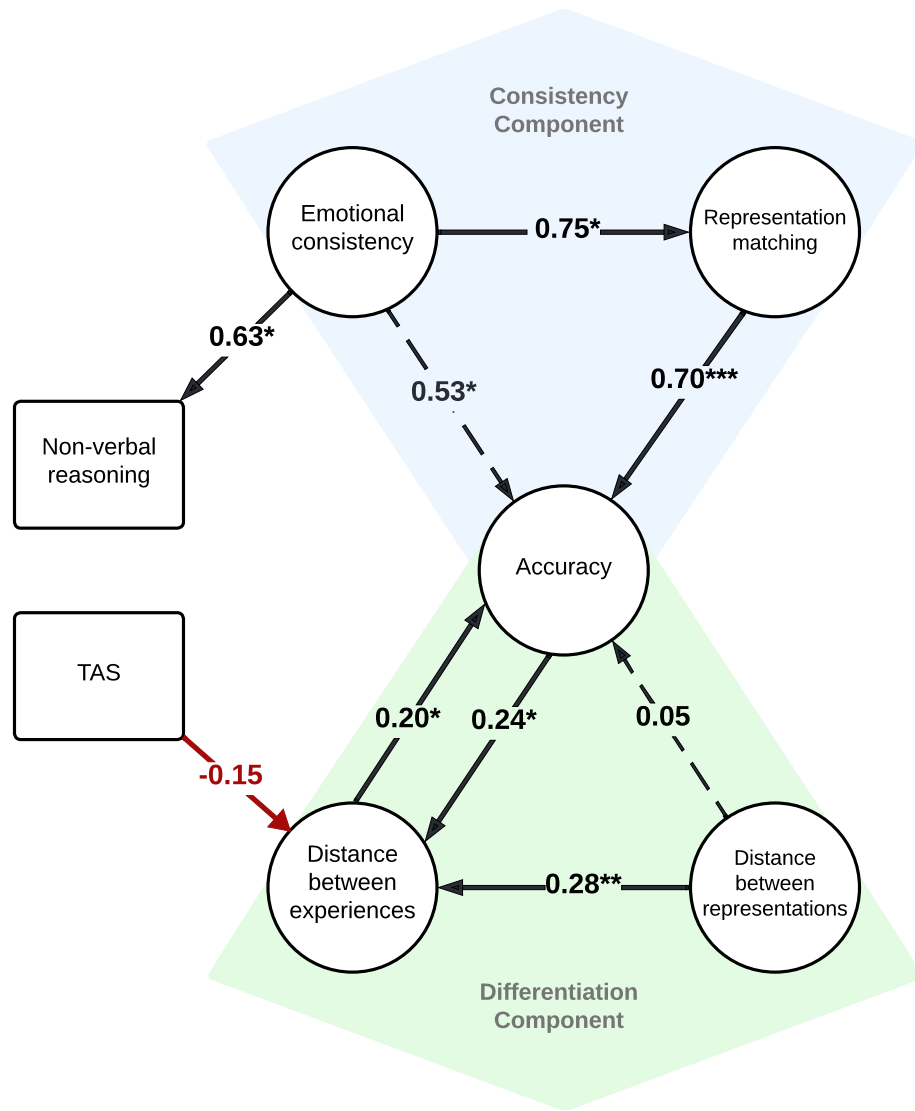
In our final structural equation model (see Fig. 6), which could account for 60.8% of the variance in emotion recognition accuracy, there were two component processes that contributed to individual differences: the *consistency component* and the *differentiation component*. With respect to the former, emotional consistency exerted an indirect effect on emotion recognition [ $z = 2.05$ ,  $b = 0.53$ ,  $p < 0.05$ ], by influencing representation matching ability [ $z = 2.06$ ,  $b = 0.75$ ,  $p < 0.05$ ], which had a direct effect on emotion recognition accuracy [ $z = 6.93$ ,  $b = 0.70$ ,  $p < 0.001$ ]. With respect to the latter, our analysis revealed that there were significant direct effects of (1) distance between emotion clusters on emotion recognition accuracy [ $z = 2.18$ ,  $b = 0.20$ ,  $p < 0.05$ ], and (2) emotion recognition accuracy on distance between emotion clusters [ $z = 2.47$ ,  $b = 0.24$ ,  $p < 0.01$ ], thus suggesting a bidirectional feedback loop between these variables. In addition, whilst distance between representations had a direct effect on distance between emotion clusters [ $z = 2.93$ ,  $b = 0.28$ ,  $p < 0.01$ ], it did not exert an indirect effect on accuracy [ $z = 1.80$ ,  $b = 0.05$ ,  $p = 0.072$ ]. Finally, our analysis also identified a significant direct effect of emotional consistency on non-verbal reasoning ability [ $z = 2.21$ ,  $b = 0.63$ ,  $p < 0.05$ ], and of alexithymia on distance between clusters [ $z = -2.27$ ,  $b = -0.15$ ,  $p < 0.05$ ] (Table 1).

## Discussion

Here we illustrated that individual differences in the experience of emotion “on the inside” are interrelated with individual differences in representations of emotional expressions, and that these sources of individual differences predict 61% of the variance in emotion recognition accuracy. In Experiment 1 we developed ( $N = 20$ ) and validated ( $N = 271$ ) “EmoMap”, a novel method to map the shape of individuals’ emotional experience landscapes. In Experiment 2 we developed “ExpressionMap” to map the landscape of ( $N = 98$ ; replication  $N = 193$ ) participants’ representations of how emotions are expressed in the outside world. Subsequently we tested for a mapping between the experience of emotion “on the inside” and representations of the way emotions are expressed in the “outside world”. Individuals with modular internal emotion maps, who had consistent and distinct emotional experiences, tended to have consistent and distinct visual representations of other people’s dynamic emotional facial expressions. Structural Equation Modelling further illustrated that such individuals tended to have correspondingly enhanced emotion recognition accuracy. Therefore, our “*Inside Out Model of Emotion Recognition*” provides new insight into the psychological mechanisms underpinning individual differences in the recognition of emotion from dynamic facial expressions.

In our final model, which explained 60.8% of the variance in emotion recognition accuracy, there were two component processes that contributed to individual differences: the *consistency component* and the *differentiation component*. Within the consistency component, which explained a larger proportion of the variance in emotion





**Figure 6.** The final structural equation model exploring the experience, visualization and recognition of emotion. Circles correspond with latent variables; rectangles correspond with manifest variables. Black arrows indicate positive relationships. Red arrows indicate negative relationships. Full line arrows correspond with direct effects; dashed line arrows correspond with indirect effects. The values displayed are the standardized beta path coefficients. The significance level for direct and indirect effects are shown by asterisks: \* $p < 0.05$ , \*\* $p \leq 0.01$ , \*\*\* $p \leq 0.001$ .

Path	Estimate	z-value	p-value	Std. b
Representation matching → accuracy	0.021	6.932	<0.001	0.700
Emotional consistency → representation matching	4.393	2.137	=0.033	0.754
Emotional consistency → representation matching → accuracy *	0.090	2.052	=0.040	0.527
Emotional consistency → NVR	0.018	2.205	=0.027	0.633
Distance between clusters → accuracy	0.039	2.179	=0.029	0.197
Accuracy → distance between clusters	1.193	2.466	=0.014	0.236
Distance between representations → distance between clusters	2.611	2.932	=0.003	0.275
Distance between representations → distance between clusters → accuracy *	0.102	1.800	=0.072	0.054
TAS → distance between clusters	-0.058	-2.269	=0.023	-0.150

**Table 1.** Parameter estimates for our final structural equation model. Standardized betas are shown in the final column (Std. b). Indirect effects are labelled with an asterisk (\*).

recognition, those with less consistent emotional experiences also had less consistent visual emotion representations, and correspondingly low emotion recognition accuracy. Interestingly, representational consistency only contributed to emotion recognition for those with a lower ability to match two visually displayed expressions. With respect to the differentiation component, having poorly differentiated representations of others' expressions, predicted poorly differentiated experiences of anger, happiness and sadness and corresponding difficulties with emotion recognition (note that this later link between experience and recognition of emotion was bi-directional). The directions of all the paths in our model were determined through systematic comparison of BIC scores. BIC comparisons revealed strong to very strong evidence that the directions in our final model were the most mathematically plausible. Nevertheless, it is important to note that structural equation modelling cannot definitively determine causality<sup>35</sup>. Thus, any directions of causality suggested by our model are merely hypotheses that should be tested via causal manipulation<sup>36</sup>.

Taking a step back from the individual path directions, it is pertinent to consider the *component processes* outlined in our final model. Although our modelling allowed for other pathways to emotion recognition difficulties—for example an emotion pathway (emotional consistency, distance between clusters) and a representation pathway (representational consistency, distance between representations)—our analyses demonstrated that the *consistency* and *differentiation* component processes were the most mathematically plausible. The emergence of these components is somewhat surprising given that EmoMap and Expression map were completed in two separate sittings (on two separate days) and that different methods were used to calculate their corresponding variables (see “Methods”). The emergence of these components *despite* their corresponding variables being calculated differently and measured across different sittings suggests that they are meaningful components of emotion recognition rather than methodological artefacts.

More generally, it is useful to consider alternative explanations for our conclusion that individual differences in emotion recognition from dynamic stimuli can, in part, be explained by individual differences in the way emotions are experienced and the way expressions are represented. A primary question concerns whether a third variable unrelated to emotion, such as participants' motivation to do well, underpins the relationships between the experience, representation and recognition of emotion. In other words, do those with more consistent experiences of emotion also have more consistent visual emotion representations and more accurate emotion recognition simply because these individuals tried harder on all tasks? Our findings suggest that this is unlikely: self-reported effort was not significantly associated with emotional consistency, representational consistency, distance between representations, matching difficulty, or representation matching, respectively (all  $p > 0.05$ ; see Supplementary Information C). In addition, although there were small-moderate correlations between effort and distance between clusters [ $R = 0.272$ ,  $p < 0.001$ ], and effort and emotion recognition accuracy [ $R = 0.236$ ,  $p = 0.001$ ] respectively, our Bonferroni-corrected partial correlations demonstrated that all the relationships we discovered remained significant after controlling for self-reported effort (see Supplementary Information C). Hence, the relationships we found between the experience, representation and recognition of emotion are not underpinned by self-reported effort. Similarly, since each of our paradigms included intricately designed attention checks, it is unlikely that differences in attention underpin the associations between these variables. Finally, one may ask whether the relationships documented here pertain specifically to the processing of emotion. That is, could it be that some individuals have consistent and distinct concepts in general and hence they are good at recognizing any complex stimuli. Our results suggest that this possibility is also unlikely: only those with consistent experiences of emotion, and not those with consistent concepts of color, had greater representational consistency and emotion recognition accuracy (see above and Supplementary Information B). Hence overall, it is unlikely that effort, attention, or another domain general process (e.g., having distinct concepts in general) underpin the associations found here. Rather our results, which have been acquired across several experiments with large samples (involving built-in replications), provide convincing evidence for links between the experience, representation and recognition of emotion.

But how do such links come about? For example, why would having inconsistent emotional experiences lead to inconsistent expression representations? As noted in our Introduction section, constructionist theories offer a theoretical framework that may help us to answer such questions. Constructionist theories of emotion (e.g.<sup>15–20</sup>) propose that children are continually constructing multimodal representations of emotions. For example, when hearing a caregiver describe a situation as “anger inducing” a child may associate their current internal sensations and prevailing visual/auditory/tactile inputs with the word “angry”. Over time “angry” ceases to be just a word and becomes a multimodal concept<sup>15</sup> and once the concept is acquired, it may function to sharpen its own conceptual boundaries<sup>14</sup>. That is, having consistent and distinct emotion concepts may help a learner focus on invariant features of facial expressions and ignore between expression variation, thus encouraging the formation of consistent (i.e., reliable) and distinct visual representations of others' expressions. Note that the reverse direction of causality is also possible: a child with consistent visual representations of others' expressions may be better able to recognize when others are angry thus providing the child with a label with which to categorize their own internal states. Such a child may have more opportunities for labelling their internal states, potentially resulting in more consistent and distinct emotional experiences. As mentioned previously, further work must test the Inside Out Model's directions of causality if we are to make more confident claims about the causal link between the experience, visual representation and recognition of emotion, and develop richer theoretical models of the developmental experiences that give rise to such links.

In addition to contributing to constructionist theories of emotion, our findings are also relevant to the face identity and signal detection literatures<sup>7–9</sup>. By demonstrating that consistent visual representations of emotional expressions facilitate the recognition of emotional expressions, we illustrate an important role for stored visual representations in *emotion* recognition, that extends beyond the known role of representations in facial *identity* recognition. Our findings are also partially in line with signal detection theory<sup>10</sup>. That is, we identified that emotion recognition was directly predicted by the consistency, but not the differentiation, of visual emotion

representations. These findings raise the possibility that there are independent contributions of these factors to emotion recognition. The lack of a significant (direct) effect of the differentiation of visual emotion representations (i.e., the distance between attributed speed) on recognition may be due to the presence of large differences in the speeds attributed to angry, happy, and sad facial expressions [angry mean (SEM) = 3.85 (0.06); happy mean (SEM) = 2.80 (0.04); sad mean (SEM) = 1.63 (0.03)], meaning that on average the representations are ‘far apart’ and instances of overlap between the signal and noise distributions are relatively uncommon. Independent of this, there may feasibly be an additional effect of the consistency of visual emotion representations (variation in attributed speeds). For example, the expectation literature would predict that more precise (i.e., a representation that is consistent in appearance across instances) representations of upcoming stimuli would precipitate increased recognition accuracy (see<sup>37–40</sup>). Future research should aim to include other emotions (e.g., surprise, disgust, and fear), likely to populate other points on the speed continuum, to identify whether this illuminates an effect of the differentiation of visual emotion representations. In the current study, we were unable to include additional emotions due to testing constraints. Including surprise, disgust and fear would have increased the duration of our test battery to over eight hours (doubling the current testing time of four hours) and compromised our ability to test such large samples (due to limits on resources). We selected high and low arousal (anger/ happiness and sadness), and positively and negatively valenced (happiness and anger/sadness), emotions to cover different regions in arousal-valence space<sup>21</sup>.

### Implications

Our Inside Out Model raises a number of testable hypotheses that may help us better understand the etiology of the emotion recognition difficulties documented in numerous clinical conditions (e.g., depression, anxiety, psychosis, eating disorders, Parkinson’s disease, and autism spectrum disorder; see<sup>41–47</sup>). In the current study, we have illuminated two component processes that may contribute to these difficulties: the differentiation component and the consistency component. With respect to the former, differences in recognizing others’ emotions may be linked to difficulties differentiating one’s own emotional states; indeed preliminary evidence supports this pathway in the context of depression, anxiety, schizophrenia, anorexia nervosa and autism (as found in<sup>25,48–52</sup>). The consistency component, on the other hand, suggests the testable hypothesis that emotion recognition difficulties in clinical conditions linked to inconsistent emotional experiences—such as bipolar disorder and psychosis, which are associated with mood fluctuations<sup>53,54</sup>—may be mediated by the (in)consistency of visual representations of emotional expressions. Identifying mechanistic pathways that explain variation in emotion recognition may help us design tailored support systems with potential impacts upon psychosocial adjustment<sup>55</sup> and psychological health and wellbeing<sup>56</sup>. Hence, future studies should aim to test these predictions.

### Limitations

The results of the current study are informative with respect to understanding the links between the experience, visual representation, and recognition of emotion from facial motion cues alone. Here, we have employed point-light displays, which provide a way of studying core dynamic cues (e.g., speed), while controlling other perceptual dimensions<sup>57,58</sup>, such as identity (e.g., gender, age, ethnicity, face attractiveness), depth, and pigmentation, which are all known to influence emotion recognition<sup>59–61</sup>. Although this allowed us to accurately assess the contribution of kinematic cues to visual emotion representations, and their subsequent effect on emotion recognition accuracy (without these other cues confounding the results), such tight control may limit the extent to which our findings generalize to full dynamic emotional expressions (e.g., full video recordings of facial expressions). It could be, for example, that the links we have demonstrated between the experience, representation and recognition of emotion exist for point-light displays, but not full emotional expressions. However, since individuals compare incoming facial expressions to stored templates, which represent the average facial expressions they have encountered previously (e.g., the average angry expression across all previous encounters<sup>62–66</sup>), it seems unlikely that the consistency of such templates would only be important when recognizing emotion from point-light displays (which are not typically encountered). Concurrently, there is no clear reason why an individual would draw on their own emotional experiences to recognize emotion, specifically in point light displays, and not in full dynamic expressions. Nevertheless, future studies are necessary to confirm whether our results generalize to full emotional expressions.

Relatedly, it is also worth noting that here we examine the consistency and differentiation of visual emotion representations specifically in the speed domain. This was an active design choice, motivated by previous evidence demonstrating the critical role of speed cues in the visual representation<sup>67</sup> and recognition<sup>22,32</sup> of emotion. Nevertheless, in future work, we will expand our paradigms to encompass other spatiotemporal emotion cues (e.g., degree of spatial exaggeration, movement onset/offset, texture, colour, etc.), thus facilitating investigations into the consistency and differentiation of visual emotion representations in other domains.

It is also important to consider the limitations of our EmoMap paradigm for assessing the experience of emotion. Although this paradigm has several methodological advantages—it can be completed online in just 25 to 35 minutes and does not require participants to translate their emotional experiences into words (see<sup>68</sup>)—there are disadvantages of using such computer-based assessments. For example, by employing images to elicit emotional reactions (as is common in the literature e.g.,<sup>6,24,26,69,70</sup>), participants may respond based on how they think they *should* feel, rather than how they truly feel. Whilst this is a possibility, we specifically address this issue in our task instructions, thus minimizing the likelihood of participants responding in this way: when describing EmoMap, we told participants that “this isn’t about what the image represents, or how you think other people, on average, respond to the images. It is about your own personal response”. Nevertheless, future investigations could benefit from employing more ecological methods such as experience sampling (e.g.,<sup>71,72</sup>), wherein participants label or rate their emotional state on several occasions throughout the day for multiple days. Using these methods,

emotion differentiation can be calculated by computing intra-class correlations, measuring consistency between emotion ratings across occasions, for each participant (see<sup>73</sup>). Such studies could then aim to test whether the ability to differentiate emotions in everyday life is associated with more differentiated visual emotion representations, and enhanced emotion recognition.

Finally, it is important to highlight the limitations of our study with respect to sample generalizability. Across both experiments discussed here, the samples were predominantly female (74.91, 46.94, and 78.76% respectively), white (58.67, 83.67, and 56.48% respectively), and from the United Kingdom (37.27, 64.29, and 41.97% respectively). Since there may be differences in the experience and recognition of emotion between males and females<sup>74–76</sup>, it may be that the results discussed here are not representative of males. Although this is possible, the evidence from our post-hoc analyses suggest that our primary effects are not moderated by sex (see Supplementary Information D). Thus, it seems that for both males and females the experience, visual representation, and recognition of emotion are all linked. Nevertheless, further work should aim to verify our results in more balanced, and/or male, samples. In addition, previous studies have found that experiences (e.g.,<sup>77–81</sup>) and visual representations (e.g.,<sup>82–85</sup>) of emotion vary across cultures. Many of these studies suggest that there may be differences specifically in the *appearance* of visual representations (e.g., individuals from Western Cultures emphasize the eyebrows and mouth in their visual representations, while those from East Asian cultures the eye region<sup>82</sup>). Although this is an important consideration, it is worth noting that, in the current study, we specifically focus on the consistency and differentiation of visual representations, rather than on the appearance of them. Since individuals across numerous cultures employ template-matching techniques (i.e., comparing incoming facial expressions to stored ‘templates’) to recognise the emotions of others<sup>62–66</sup>, it seems unlikely that the *consistency* of such templates would be important in one culture but not another. Nevertheless, future studies should aim to the Inside Out Model across different cultures.

## Methods

### Experiment 1: developing EmoMap

This study was approved by the Science, Technology, Engineering and Mathematics (STEM) ethics committee at the University of Birmingham (ERN\_16-0281AP9D) and was conducted in accordance with the principles of the revised Helsinki Declaration. Informed consent was obtained from all participants.

#### Participants

271 participants were recruited via the School of Psychology’s Research Participation Scheme database and Prolific. Individuals were eligible to take part in the study if they were between the ages of 18 and 65, fluent in English, had normal or corrected-to-normal vision, and had access to a computer/laptop with an internet connection. The sample size was based on an a priori power analysis conducted using G\*power<sup>86</sup>. To replicate the association between alexithymia and emotion differentiation in Erbas et al.,<sup>26</sup> 97 participants were required to have 95% power at alpha level 0.05. However, since effect sizes are commonly inflated<sup>87</sup>, and we were utilizing a more complicated analysis (a linear mixed effects model in which we control for the other relevant demographic variables known to be associated with the experience and perception of emotion), we recruited a larger number of participants (N = 271); almost triple the sample size generated in our power calculation.

Participant characteristics are displayed in Table 2. Information regarding participants’ ethnicities is reported in Supplementary Information E. Notably, four participants in the sample (1.48%) reported that they had a diagnosis of autism spectrum disorder. Therefore, we conducted our analyses twice, first including these participants and then excluding them. Since the general pattern of results was unaffected by their removal, we included these participants in our final statistical analyses.

#### Procedures

Participants completed demographics questions, followed by the Autism Quotient<sup>88</sup> and Toronto Alexithymia Scale<sup>89</sup> on Qualtrics (<https://www.qualtrics.com>). Subsequently, participants completed our EmoMap paradigm (openly available at <https://app.gorilla.sc/openmaterials/447800>) followed by the Matrix Reasoning Item Bank<sup>90</sup> on Gorilla (<https://gorilla.sc>). All participants completed the study online.

#### Materials and stimuli

**Demographics: questionnaires and tasks.** The level of autistic traits of participants was assessed via the Autism Quotient<sup>88</sup>. The AQ is a 50-item questionnaire scored from 0 to 50, with higher scores representing higher levels

Variable	Participants (N = 271)
Sex	68 male, 203 female
Age	24.00 (9.16)
NVR	60.22% (15.35%)
AQ-50	19.11 (6.85)
TAS-20	48.17(12.08)

**Table 2.** Means and standard deviations of participant characteristics. In the column on the right-hand side, means are followed by standard deviation in parentheses.

of autistic characteristics. The level of alexithymic traits of participants was measured via the 20-item Toronto Alexithymia Scale<sup>89</sup>. The TAS comprises 20 items rated on a five-point Likert scale, ranging from 1, strongly disagree, to 5, strongly agree. Total scores on the scale range from 20 to 100, with higher scores indicating higher levels of alexithymic traits. Non-verbal reasoning ability was assessed via the Matrix Reasoning Item bank (MaRs-IB<sup>90</sup>). Each item in the MaRs-IB consists of a 3 × 3 matrix. In this matrix, eight of the nine available cells are filled with abstract shapes, and one cell in the bottom right-hand corner is left empty. Participants are required to complete the matrix by selecting the missing shape from four possible options. To correctly identify the answer, participants have to deduce relationships between the shapes in the matrix (which vary in shape, color, size and position). Non-verbal reasoning scores represent the percentage of trials in which participants provide the correct response.

**EmoMap task.** There were two key parts of the EmoMap task. In the first part, on each trial participants viewed pairs of images from the Nencki Affective Picture System<sup>29</sup>, and were instructed to “think about what feelings arise when you look at each of these images. Now please rate how SIMILAR those two feelings are”. Participants made their ratings on a visual analogue scale (with a step size of 0.0001) ranging from 0, ‘Not at all similar’ to 10, ‘Very similar’. An advantage of the EmoMap paradigm is that it provides a tool to measure emotion differentiation without requiring participants to produce emotion labels, unlike existing tasks (see<sup>68,91</sup> for a full discussion).

The chosen images were known to be effective at selectively inducing anger, happiness or sadness across large samples (N = 124;<sup>30</sup>), and generated distinct emotion clusters based on graph theory analyses with pilot study data (see Supplementary Information F). In this task, there were five images for each emotion (anger, happiness and sadness) resulting in 15 different images and 105 unique image combinations (and therefore 105 trials): 30 within emotion-category combinations (10 for anger, 10 for happiness and 10 for sadness) and 75 between emotion-category combinations (25 angry-sad, 25 angry-happy, 25 happy-sad). A reaction time check was incorporated to prevent participants responding too quickly (i.e., without thinking). Responses faster than 1000 ms resulted in an error message (“Too Fast. Our algorithm has detected that you might need to take longer to think through your answer. You will now incur a 5 second penalty and then will be asked to do the trial again”) and a 5 second penalty, and then the trial was re-started. This threshold was selected to give participants sufficient time to view the images, detect and compare the emotions evoked by each of them, and then respond by clicking on the scale.

To map the shape and size of participants’ internal emotional landscapes, similarity ratings were transformed into (Euclidean) distance scores through multidimensional scaling (using the Scikit-learn library in Python). Multidimensional scaling (MDS) is a statistical technique that represents objects (emotional images, lexical items etc.) as points in multidimensional space wherein close similarity between objects corresponds to close distances between the corresponding points in the representation<sup>92</sup>. The distance between points in multidimensional space can then be plotted (see Fig. 1). Mean distances within specific emotion clusters comprised the average of the Euclidean distances for the 10 angry-angry, 10 happy-happy and 10 sad-sad image pairs, respectively. Mean distances between specific emotion clusters comprised the mean of the Euclidean distances for the 25 angry-happy, 25 angry-sad, and 25 happy-sad image pairs, respectively. We then computed mean distances within and between clusters by averaging across emotions/emotion pairs. Larger distances between and within emotion clusters reflect greater emotion differentiation.

The second part of our EmoMap paradigm was inspired by the work of Huggins and colleagues<sup>93</sup>. In this part of the task, on each trial, participants were presented with three images from the Nencki Affective Picture System<sup>29</sup>, and were required to make a decision. This task involved four conditions: one non-emotional control condition, and three emotional experimental conditions exploring the experience of anger, happiness and sadness respectively. First, participants completed the non-emotional control condition. In this condition, participants were required to select which of the three (neutral) images they found most colorful using their mouse cursor. Two of these images were in color and one was in grayscale, thus serving as an attention check. If participants selected the grayscale image, they were presented with an error message, incurred a 5 second penalty, and then were asked to do the trial again. Following this, participants completed the three experimental conditions in a random order. In these conditions, participants were required to select which of the three images made them feel most angry, happy or sad (e.g., in the angry condition, participants had to decide which of the two images made them most angry) using their mouse cursor. As was the case in the control condition, there was a ‘trap’ image on each trial in the emotional conditions. On each trial, two of the images were strong inducers of the target emotion (e.g., sadness), and one was a strong inducer of another emotion (e.g., happiness), thus serving as an attention check. If participants selected the image that strongly induced the non-target emotion, they were presented with the error message discussed above, incurred a 5second penalty, and then were asked to do the trial again. Within each condition, there were 11 target (i.e., non-trap) images which were presented in all possible unique pairs across 55 trials. The images that were selected had previously been identified as successful inducers of the target emotion<sup>30</sup>. In addition, in order to make the experimental conditions comparable, we ensured that the mean intensity ratings (angry = 3.53; happy = 3.50; sad = 3.56) and standard deviation of intensity ratings across images within a condition (angry = 0.80; happy = 0.80; sad = 0.81) were similar for each emotion.

Consistency scores were calculated for each condition in line with the logical consistency of a participants’ decisions. To illustrate, if a participant selects image A over image B ( $A > B$ ) and image B over image C ( $B > C$ ), these decisions are all consistent with one another. However, if the participant then selected image C over image A, this would be inconsistent with their previous judgments. Consistency requires participants to differentiate precisely between the intensity of emotion evoked by each image. Thus, inconsistent decisions are likely to stem from inconsistencies in how individuals experience an emotion across multiple instances.

We followed the procedures of Huggins et al.,<sup>93</sup> to calculate consistency. We first quantified each participant’s image rankings by summing the number of times they chose each image. If a participant made completely



consistent decisions in a set, rank scores would follow a linear sequence: the image they found most emotionally intense (or colorful) should be chosen in all ten trials it appeared (score = 10), the second-highest should be chosen in nine of ten trials (score = 9), and so on. The image they found least emotionally intense (or colorful) should never be chosen (score = 0). Subsequently, we examined how image rankings related to the decisions made on each trial. Since images with higher rank scores should elicit stronger emotional responses than those with lower rank scores, an inconsistent decision would occur when a lower-ranking image is chosen over a higher-ranking image. For each trial, the rank score of the unchosen item was subtracted from the rank score for the chosen item, producing item differences. For inconsistent decisions the item difference would be less than or equal to zero. More severe inconsistencies (e.g. choosing the lowest ranked image over the highest ranked image) result in more negative item differences. Item differences were then summed, per condition, to produce total consistency scores, with greater scores reflecting higher consistency. If a participant made no inconsistent decisions within a condition, their score would be 220.

## Experiment 2: developing ExpressionMap

### Participants

The first (“original”) sample for Experiment 2 comprised 98 participants recruited via Prolific. The second, replication, sample comprised 193 participants recruited via the School of Psychology’s Research Participation Scheme database and Prolific. For both samples, individuals were eligible to take part if they were between the ages of 18 and 65, fluent in English, had normal or corrected-to-normal vision, and had access to a computer/laptop with Google Chrome and an internet connection. The sample size for our replication sample was based on an a priori power calculation using GLIMMPSE<sup>94</sup>. To have 90% power to replicate our finding from sample one that representational consistency predicted emotion recognition accuracy, 68 participants were necessary (alpha level = 0.05). Since effect sizes are commonly inflated<sup>87</sup> and using larger samples improves the precision of parameter estimates<sup>95</sup>, we recruited a larger number of participants (N = 193; almost triple the sample size generated in our power calculation).

Participant characteristics are displayed in Table 3. Information regarding participants’ ethnicities is reported in Supplementary Information E. Notably, one participant in the original sample (1.02%) and two participants in the replication sample (1.02%) reported that they had a diagnosis of autism spectrum disorder. Therefore, we conducted our analyses both including, and then excluding, these participants. Since the general pattern of results was unaffected by their removal, we included these participants in our final statistical analyses.

### Procedures

First, informed consent was obtained from all participants before conducting the study. Participants in the original sample completed demographics questions, followed by the 50-item Autism Quotient<sup>88</sup>, and the 20-item Toronto Alexithymia Scale<sup>89</sup> on Qualtrics (<https://www.qualtrics.com>). Following this, these participants completed three tasks that employed dynamic point light displays (a series of dots that convey biological motion) of angry, happy and sad facial expressions (PLFs) on Gorilla (<https://gorilla.sc>). Participants completed ExpressionMap followed by the Visual Matching task, followed by the PLF Emotion Recognition task (openly available at <https://app.gorilla.sc/openmaterials/447800>). Finally, participants completed the MaRs-IB<sup>90</sup>. For those in the replication sample, participation was split across two parts. In part one, participants completed demographics questions, the AQ, TAS and EmoMap paradigm. In part two, which was completed in a separate sitting at least 24 hours after finishing part one, participants completed ExpressionMap, the Visual Matching Task, the PLF Emotion Recognition task and the MaRs-IB. For both samples, all parts of the study were completed online.

### Materials and stimuli

**ExpressionMap.** In this task, on each trial, participants were presented with a dynamic point light display of the face (PLF; on average approximately 6 seconds in length) that looped such that it played continuously. Participants were instructed to “move the dial to change the speed of this video until it matches the speed of a typical ANGRY/HAPPY/SAD expression”.

The PLFs were originally created by asking actors to read a sentence (“my name is John and I’m a scientist”) in a happy, angry or sad manner<sup>32</sup>. The emotion depicted in the stimulus video matched the instructed emotion (i.e., on a trial where an angry facial expression was presented, participants were only asked to manipulate the speed of the video so that it matched the speed of a typical angry expression). Consequently, participants were matching the speed of the displayed PLF to their imagined visual representation of that expression (the

Variable	Experiment 2, original sample (n = 98)	Replication sample (n = 193)
Sex	52 male, 46 female	41 male, 152 female
Age	33.34 (9.79)	23.41 (9.04)
NVR	58.45% (16.62%)	61.24% (14.79%)
AQ-50	18.65 (7.64)	18.94 (6.79)
TAS-20	46.00 (11.82)	48.13 (11.58)

**Table 3.** Means and standard deviations of participant characteristics. In the column on the right-hand side, means are followed by standard deviation in parentheses.

speed they would imagine “in their mind’s eye”). Participants could change the speed of the video by moving a dial clockwise to increase the speed of the animation or anticlockwise to decrease the speed of the animation. The minimum and maximum point on the dial corresponded with 25% and 300% of the recorded speed respectively. Once participants were satisfied, they pressed the spacebar to continue. There was no time limit for participants to respond on each trial. Participants were shown four repetitions of each PLF stimulus video (each one starting at a random speed) across four actors. This resulted in 16 videos per emotion (4 actors  $\times$  4 repetitions  $\times$  3 emotions = 48 trial in total). Participants completed three practice trials (one for each emotion at 100% starting speed) and then the 48 randomly ordered experimental trials across three blocks. Participants were encouraged to take breaks between blocks.

The ExpressionMap task was adapted from Keating et al.<sup>67</sup>. In the current study we improved the task by (a) using a dial, instead of the slider used previously, thus making the minimum and the maximum points on the scale more ambiguous, (b) starting each video at a random speed thus reducing potential response biasing, (c) setting the initial dial position to a random orientation that did not correspond to starting speed, thus ensuring that the minimum and maximum points, and the point of the 100% recorded speed were at different spatial locations on the dial—as a result, participants were unable to be consistent simply by selecting a similar location on the dial each time—, (d) incorporating a reaction time check—when participants responded faster than 5 seconds on a trial, they were presented with an error message, incurred a 5 second penalty, and then were asked to do the trial again and, (e) incorporating a walk-through video to facilitate comprehension of task instructions.

Whereas existing methods aim to construct comprehensive representations of emotional expressions (e.g.<sup>83,85,96</sup>), ExpressionMap seeks to assess accompanying features of those representations (e.g., speed, consistency and differentiation; see<sup>97</sup>). ExpressionMap provides an index of the percentage speed attributed to each of the stimulus videos by participants (e.g., if a participant attributes 130% speed to an expression, their representation for that expression is 1.3 times faster than the recorded speed). Following the procedures outlined in Keating et al.<sup>67</sup>, we calculated the true speed attributed to each of the PLFs (in pixels per frame) by multiplying the percentage speed attributed, divided by 100, with the speed of the actor’s facial movement in the original video. For example, for a trial in which a participant attributed 200% speed to a face moving at 2.5 pixels/frame, the true speed attributed to the expression would be 5 pixels/frame [i.e.,  $(200 \div 100) \times 2.5$ ] (see<sup>67</sup> for more information).

This task operates on the premise that, compared to participants with consistent visual representations, those with less consistent representations of emotion would attribute more variable speeds to the expressions<sup>98</sup>. For instance, someone with a consistent visual representation of anger would attribute similar speeds across repetitions (e.g., by attributing 120% speed, 121% speed and 119% speed to the angry expression). In contrast, someone with a less consistent visual representation would be more variable (e.g., by attributing 120% speed, 60% speed and 180% speed to an angry expression). Therefore, to index the consistency of visual emotion representations, we took the standard deviation of the speeds attributed to one emotion for one actor (i.e., actor 1, angry expression) across the 4 video repetitions. Following this, we multiplied standard deviation scores by -1 so that our variable would now represent consistency (note that in Figs. 4 and 5 we also added a constant of 2.52, since the lowest score was -2.52, to facilitate interpretation). We then calculated mean representational consistency for each of the emotions (angry, happy and sad) by taking a mean of the consistency scores for each actor within an emotion (e.g., taking a mean of the consistency scores for angry expressions across actors 1, 2, 3 and 4). Mean representational consistency was calculated by taking a mean of the consistency scores for the angry, happy and sad PLFs.

Finally, this task also provides an index of the ‘distance’ between emotions in participants’ visual representations of facial expressions. To calculate distance scores, we subtracted the speed attributed to one emotion from the speed attributed to another and then took the absolute value of this number. For example, to calculate the distance between happy and angry PLFs, we subtracted the mean speed attributed to happy PLFs from the mean speed attributed angry PLFs, and then took the absolute value. Mean distance was calculated by taking a mean of the scores for the angry-happy, angry-sad, and happy-sad distances.

**Visual matching task.** We reasoned that an individual might have beautifully consistent mental representations of others’ expressions and still struggle to recognize others’ emotions due to an inability to match the incoming expression data with the stored representation. Thus, we developed the Visual Matching task to assess how well participants can visually match the speed of one expression to another (displayed) expression. Each trial began with a PLF stimulus video on the left-hand side of the screen. After this video had played once, the same PLF stimulus video also appeared on the right-hand side of screen, moving at a random speed, and continued to play in a loop. Participants were instructed to “move the dial to change the speed of the video on the right until it matches the speed of the video on the left”. Turning the dial clockwise increased speed. Turning the dial anticlockwise decreased speed. The minimum and maximum points on the dial corresponded with 25% speed and 300% of the recorded speed respectively (participants were not explicitly informed about this). Once the participant was satisfied, they pressed spacebar to continue. Participants were shown four repetitions of each PLF stimulus video for each of the four actors; each repetition had a different starting speed. In each full set of 16 (4 actors  $\times$  4 repetitions) stimulus videos for an emotion, the starting speed ranged from 50 to 200% of the recorded speed, in 10% increments (i.e., 50%, 60%, 70%, 80%, 90%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%). This range of starting speeds ensured that participants were able to match across a variety of speeds. Participants completed three practice trials (one for each emotion at 100% starting speed) and then the 48 randomly ordered experimental trials across three blocks. Participants were given the opportunity to take breaks between blocks.

The Visual Matching task provides an index of how well participants can visually match the speed of one expression to another. To calculate deviation scores, we subtracted the percentage speed attributed to the

expression on the right from the percentage speed attributed to the expression on the left, and then took the absolute value. As such, deviation scores provide a measure of how far away the speeds of the two animations were (irrespective of whether they attributed too high or too low speed). Finally, we calculated mean deviation scores by taking a mean of all of the absolute deviation scores. Higher deviation scores represent greater difficulties matching the two expressions.

**PLF emotion recognition task.** In the PLF Emotion Recognition task<sup>32</sup>, participants viewed dynamic PLFs, created from videos of four actors verbalizing sentences whilst posing three target emotions (angry, happy and sad). PLFs were adapted (see<sup>32</sup> for further detail) to achieve three spatial movement levels, ranging from decreased to increased spatial movement (50%, 100% and 150% spatial movement), and three kinematic levels, ranging from reduced to increased speed (50%, 100% and 150% original stimulus speed). Each trial in this task began with the presentation of a (silent) PLF video displaying one of the three emotions, at one of the three spatial and three kinematic levels. After watching the video, participants were required to rate how angry, happy and sad the person was feeling on three visual analogue scales (presented in a random order) ranging from ‘Not at all angry/happy/sad’ to ‘Very angry/happy/sad’. Each trial lasted approximately 25 seconds. Participants completed three practice trials and then 108 randomly ordered experimental trials (12 per condition) across three blocks. Participants were encouraged to take a break between blocks.

The three emotion rating responses for each trial were transformed into magnitude scores from zero to ten (with zero representing a response of ‘Not at all’ and ten representing ‘Very’) to three decimal places. Emotion recognition accuracy scores were calculated by subtracting the mean of the two incorrect emotion ratings from the correct emotion rating. For example, for a trial in which a happy PLF was displayed, the mean rating of the two incorrect emotions (angry and sad) was subtracted from the rating for the correct emotion (happy). Mean emotion recognition accuracy was calculated by taking a mean of accuracy scores across all emotions and levels of the spatial and kinematic manipulation.

#### Transparency and openness

In this manuscript, we report how we determined our sample sizes, all data exclusions, all manipulations, and how we calculated all measures. All datafiles, data-processing code, analysis scripts, and tasks are openly available at <https://osf.io/hd8u2/wiki/home/>. The data were processed and analyzed using R (R Studio version 2021.09.2), Python (Jupyter Notebook version 6.4.8), and JASP (version 0.16). All our linear mixed effects models were conducted in R Studio using the lmer function (from the *lme4* package). For these models, we also used the Anova function (from the *car* package) to conduct a Type III ANOVA with a Kenward-Roger<sup>99</sup> approximation for degrees of freedom, as supported by Luke<sup>100</sup>. For all linear mixed effects models, the relationships between the experience, representation, and recognition of emotion hold when the control variables are included (as reported in the Results section) and excluded, thus affording us greater confidence in our findings. In R Studio, we also conducted a random forest analysis<sup>33</sup> employing the Boruta wrapper algorithm (Boruta function from *Boruta* package,<sup>34</sup>) and structural equation modelling using the sem() function (from the *lavaan* package). In JASP, we conducted Bayesian linear regressions to determine the relative strength of evidence for the experimental versus null hypotheses. For these analyses, we followed the classification scheme from Lee and Wagenmakers<sup>101</sup>: BF<sub>10</sub> values between one and three reflect weak evidence, between three and ten moderate evidence, between ten and 100 strong evidence, and greater than 100 extreme evidence for the *experimental* hypothesis.

#### Data availability

All datafiles, data-processing code, analysis scripts, and tasks are openly available at <https://osf.io/hd8u2/wiki/home/>.

Received: 19 May 2023; Accepted: 27 November 2023

Published online: 06 December 2023

#### References

- Argaud, S., V erin, M., Sauleau, P. & Grandjean, D. Facial emotion recognition in Parkinson's disease: A review and new hypotheses. *Mov. Disord.* **33**(4), 554–567. <https://doi.org/10.1002/mds.27305> (2018).
- Hofer, A. *et al.* Facial emotion recognition and its relationship to symptomatic, subjective, and functional outcomes in outpatients with chronic schizophrenia. *Eur. Psychiatry* **24**(1), 27–32. <https://doi.org/10.1016/j.eurpsy.2008.06.008> (2009).
- Hoertnagl, C. M. *et al.* Facial emotion recognition and its relationship to subjective and functional outcomes in remitted patients with bipolar I disorder. *Bipolar Disord.* **13**(5–6), 537–544. <https://doi.org/10.1111/j.1399-5618.2011.00947.x> (2011).
- Huang, Y. A. *et al.* Studying emotion theories through connectivity analysis: Evidence from generalized psychophysiological interactions and graph theory. *Neuroimage* **172**, 250–262. <https://doi.org/10.1016/j.neuroimage.2018.01.026> (2018).
- Stolier, R. M., Hehman, E., Keller, M. D., Walker, M. & Freeman, J. B. The conceptual structure of face impressions. *Proc. Natl. Acad. Sci.* **115**(37), 9210–9215. <https://doi.org/10.1073/pnas.1807222115> (2018).
- Israelashvili, J., Oosterwijk, S., Sauter, D. & Fischer, A. Knowing me, knowing you: Emotion differentiation in oneself is associated with recognition of others' emotions. *Cogn. Emot.* **33**(7), 1461–1471. <https://doi.org/10.1080/02699931.2019.1577221> (2019).
- Bruce, V. & Young, A. Understanding face recognition. *Br. J. Psychol.* **77**(3), 305–327. <https://doi.org/10.1111/j.2044-8295.1986.tb02199.x> (1986).
- Young, A. W. & Bruce, V. Understanding person perception. *Br. J. Psychol.* **102**(4), 959–974. <https://doi.org/10.1111/j.2044-8295.2011.02045.x> (2011).
- Etchells, D. B., Brooks, J. L. & Johnston, R. A. Evidence for view-invariant face recognition units in unfamiliar face learning. *Q. J. Exp. Psychol.* **70**(5), 874–889. <https://doi.org/10.1080/17470218.2016.1248453> (2017).
- McNicol, D. *A primer of signal detection theory* (Psychology Press, 2005).
- Fessler, P. K., Lenorovitz, D. R. & Yoblick, D. A. Time delay and similarity effects in facial recognition. *J. Appl. Psychol.* **59**(4), 490. <https://doi.org/10.1037/h0037193> (1974).

12. Gentner, D. Bootstrapping the mind: Analogical processes and symbol systems. *Cogn. Sci.* **34**(5), 752–775. <https://doi.org/10.1111/j.1551-6709.2010.01114.x> (2010).
13. Gentner, D. & Medina, J. Similarity and the development of rules. *Cognition* **65**(2–3), 263–297. [https://doi.org/10.1016/S0010-0277\(98\)00002-X](https://doi.org/10.1016/S0010-0277(98)00002-X) (1998).
14. Boroditsky, L. Comparison and the development of knowledge. *Cognition* **102**(1), 118–128. <https://doi.org/10.1016/j.cognition.2002.08.001> (2007).
15. Hoemann, K., Xu, F. & Barrett, L. F. Emotion words, emotion concepts, and emotional development in children: A constructionist hypothesis. *Dev. Psychol.* **55**(9), 1830. <https://doi.org/10.1037/dev0000686> (2019).
16. Barrett, L. F. Solving the emotion paradox: Categorization and the experience of emotion. *Pers. Soc. Psychol. Rev.* **10**(1), 20–46. [https://doi.org/10.1207/s15327957pspr1001\\_2](https://doi.org/10.1207/s15327957pspr1001_2) (2006).
17. Nook, E. C., Lindquist, K. A. & Zaki, J. A new look at emotion perception: Concepts speed and shape facial emotion recognition. *Emotion* **15**(5), 569. <https://doi.org/10.1037/a0039166> (2015).
18. Satpute, A. B. *et al.* Emotions in “black and white” or shades of gray? How we think about emotion shapes our perception and neural representation of emotion. *Psychol. Sci.* **27**(11), 1428–1442. <https://doi.org/10.1177/0956797616661555> (2016).
19. Lindquist, K. A. & Barrett, L. F. Constructing emotion: The experience of fear as a conceptual act. *Psychol. Sci.* **19**(9), 898–903 (2008).
20. Lindquist, K. A., Satpute, A. B. & Gendron, M. Does language do more than communicate emotion?. *Curr. Direct. Psychol. Sci.* **24**(2), 99–108. <https://doi.org/10.1177/0963721414553440> (2015).
21. Russell, J. A. A circumplex model of affect. *J. Pers. Soc. Psychol.* **39**(6), 1161. <https://doi.org/10.1037/h0077714> (1980).
22. Keating, C. T., Fraser, D. S., Sowden, S. & Cook, J. L. Differences between autistic and non-autistic adults in the recognition of anger from facial motion remain after controlling for alexithymia. *J. Autism Dev. Disord.* **52**(4), 1855–1871. <https://doi.org/10.1007/s10803-021-05083-9> (2022).
23. Keating, C. T., Ichijo, E. & Cook, J. L. Autistic adults exhibit highly precise representations of others’ emotions but a reduced influence of emotion representations on emotion recognition accuracy. *Sci. Rep.* **13**(1), 11875. <https://doi.org/10.1038/s41598-023-39070-0> (2023).
24. Keating, C. T., Kraaijkamp, C., & Cook, J. (pre-print). Similarities and differences in the psychological mechanisms involved in autistic and non-autistic emotion recognition. *PsyArxiv*.
25. Erbas, Y., Ceulemans, E., Boonen, J., Noens, I. & Kuppens, P. Emotion differentiation in autism spectrum disorder. *Res. Autism Spectr. Disord.* **7**(10), 1221–1227. <https://doi.org/10.1016/j.rasd.2013.07.007> (2013).
26. Erbas, Y., Ceulemans, E., Lee Pe, M., Koval, P. & Kuppens, P. Negative emotion differentiation: Its personality and well-being correlates and a comparison of different assessment methods. *Cogn. Emotion* **28**(7), 1196–1213. <https://doi.org/10.1080/0269931.2013.875890> (2014).
27. Cook, R., Brewer, R., Shah, P. & Bird, G. Alexithymia, not autism, predicts poor recognition of emotional facial expressions. *Psychol. Sci.* **24**(5), 723–732. <https://doi.org/10.1177/0956797612463582> (2013).
28. Milosavljevic, B. *et al.* Alexithymia in adolescents with autism spectrum disorder: Its relationship to internalising difficulties, sensory modulation and social cognition. *J. Aut. Dev. Disord.* **46**(4), 1354–1367. <https://doi.org/10.1007/s10803-015-2670-8> (2016).
29. Marchewka, A., Żurawski, L., Jednoróg, K. & Grabowska, A. The Nencki affective picture system (NAPS): Introduction to a novel, standardized, wide-range, high-quality, realistic picture database. *Behav. Res. Methods* **46**(2), 596–610. <https://doi.org/10.3758/s13428-013-0379-1> (2014).
30. Riegel, M. *et al.* Characterization of the Nencki affective picture system by discrete emotional categories (NAPS BE). *Behav. Res. Methods* **48**(2), 600–612. <https://doi.org/10.3758/s13428-015-0620-1> (2016).
31. Nemiah, J. C., Freyberger, H. & Sifneos, P. E. Alexithymia: A view of the psychosomatic process. *Mod. Trends Psychos. Med.* **3**, 430–439 (1976).
32. Sowden, S., Schuster, B. A., Keating, C. T., Fraser, D. S. & Cook, J. L. The role of movement kinematics in facial emotion expression production and recognition. *Emotion*. **21**(5), 1041–1061. <https://doi.org/10.1037/emo0000835> (2021).
33. Breiman, L. Random forests. *Mach. Learn.* **45**(1), 5–32. <https://doi.org/10.1023/A:1010933404324> (2001).
34. Kursa, M. B. & Rudnicki, W. R. Feature selection with the Boruta package. *J. Stat. Softw.* **36**, 1–13. <https://doi.org/10.18637/jss.v036.i11> (2010).
35. Bollen, K. A., & Pearl, J. Eight myths about causality and structural equation models. In *Handbook of causal analysis for social research* (pp. 301–328) (Springer, Dordrecht, 2013).
36. Happé, F., Cook, J. L. & Bird, G. The structure of social cognition: In (ter) dependence of sociocognitive processes. *Annu. Rev. Psychol.* **68**(1), 243–267. <https://doi.org/10.1146/annurev-psych-010416-044046> (2017).
37. Clark, A. *Surfing uncertainty: Prediction, action, and the embodied mind* (Oxford University Press, 2015).
38. Friston, K. The free-energy principle: A unified brain theory?. *Nat. Rev. Neurosci.* **11**(2), 127–138. <https://doi.org/10.1038/nrn2787> (2010).
39. Hohwy, J. *The predictive mind* (Oxford University Press, 2013).
40. Sevi, L. *et al.* Egocentric biases are predicted by the precision of self-related predictions. *Cortex* <https://doi.org/10.1016/j.cortex.2022.04.021> (2022).
41. Dalili, M. N., Penton-Voak, I. S., Harmer, C. J. & Munafò, M. R. Meta-analysis of emotion recognition deficits in major depressive disorder. *Psychol. Med.* **45**(6), 1135–1144. <https://doi.org/10.1017/S0033291714002591> (2015).
42. Demenescu, L. R., Kortekaas, R., den Boer, J. A. & Aleman, A. Impaired attribution of emotion to facial expressions in anxiety and major depression. *PLoS One* **5**(12), e15058. <https://doi.org/10.1371/journal.pone.0015058> (2010).
43. Kohler, C. G., Walker, J. B., Martin, E. A., Healey, K. M. & Moberg, P. J. Facial emotion perception in schizophrenia: A meta-analytic review. *Schizophrenia Bull.* **36**(5), 1009–1019. <https://doi.org/10.1093/schbul/sbn192> (2010).
44. Preti, A., Siddi, S., Marzola, E., & Abbate Daga, G. Affective cognition in eating disorders: a systematic review and meta-analysis of the performance on the “Reading the Mind in the Eyes” Test. *Eat. Weight Disord. Stud. Anorexia Bulimia Obes.* 1–17. <https://doi.org/10.1007/s40519-022-01393-8> (2022).
45. Gray, H. M. & Tickle-Degnen, L. A meta-analysis of performance on emotion recognition tasks in Parkinson’s disease. *Neuropsychology* **24**(2), 176. <https://doi.org/10.1037/a0018104> (2010).
46. Keating, C. T. & Cook, J. L. Facial expression production and recognition in autism spectrum disorders: A shifting landscape. *Child Adolesc. Psychiatr. Clin.* **29**(3), 557–571. <https://doi.org/10.1016/j.chc.2020.02.006> (2020).
47. Nook, E. C. Emotion differentiation and youth mental health: Current understanding and open questions. *Front. Psychol.* **32**, 81. <https://doi.org/10.3389/fpsyg.2021.700298> (2021).
48. Demiralp, E. *et al.* Feeling blue or turquoise? Emotional differentiation in major depressive disorder. *Psychol. Sci.* **23**(11), 1410–1416. <https://doi.org/10.1177/0956797612444903> (2012).
49. Kashdan, T. B. & Farmer, A. S. Differentiating emotions across contexts: Comparing adults with and without social anxiety disorder using random, social interaction, and daily experience sampling. *Emotion* **14**(3), 629. <https://doi.org/10.1037/a0035796> (2014).
50. Kimhy, D. *et al.* Emotional granularity and social functioning in individuals with schizophrenia: An experience sampling study. *J. Psychiatr. Res.* **53**, 141–148. <https://doi.org/10.1016/j.jpsychires.2014.01.020> (2014).



51. Selby, E. A. *et al.* Nothing tastes as good as thin feels: Low positive emotion differentiation and weight-loss activities in anorexia nervosa. *Clin. Psychol. Sci.* **2**(4), 514–531. <https://doi.org/10.1177/2167702613512794> (2014).
52. Williams-Kerver, G. A. & Crowther, J. H. Emotion differentiation and disordered eating behaviors: The role of appearance schemas. *Eat. Behav.* **37**, 101369. <https://doi.org/10.1016/j.eatbeh.2020.101369> (2020).
53. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (5th ed.). *Am. J. Psychiatry* (2013).
54. Marwaha, S., Broome, M. R., Bebbington, P. E., Kuipers, E. & Freeman, D. Mood instability and psychosis: Analyses of British national survey data. *Schizophrenia Bull.* **40**(2), 269–277. <https://doi.org/10.1093/schbul/sbt149> (2014).
55. Hall, J. A., Andrzejewski, S. A. & Yopchick, J. E. Psychosocial correlates of interpersonal sensitivity: A meta-analysis. *J. Nonverbal Behav.* **33**(3), 149–180. <https://doi.org/10.1007/s10919-009-0070-5> (2009).
56. Bänziger, T. Accuracy of judging emotions. *The social psychology of perceiving others accurately*, 23–51. <https://doi.org/10.1017/CBO9781316181959.002> (2016).
57. Bidet-Ildei, C., Decatoire, A. & Gil, S. Recognition of emotions from facial point-light displays. *Front. Psychol.* **11**, 1062. <https://doi.org/10.3389/fpsyg.2020.01062> (2020).
58. Pavlova, M. A. Biological motion processing as a hallmark of social cognition. *Cerebral Cortex* **22**(5), 981–995. <https://doi.org/10.1093/cercor/bhr156> (2012).
59. Jelili, S. *et al.* Impaired recognition of static and dynamic facial emotions in children with autism spectrum disorder using stimuli of varying intensities, different genders, and age ranges faces. *Front. Psychiatry* **12**, 693310. <https://doi.org/10.3389/fpsyg.2021.693310> (2021).
60. Wang, L., Chen, W. & Li, H. Use of 3D faces facilitates facial expression recognition in children. *Sci. Rep.* **7**(1), 45464. <https://doi.org/10.1038/srep45464> (2017).
61. Yasuda, M. *Color and facial expressions* (University of Nevada, 2007).
62. Minemoto, K., Ueda, Y. & Yoshikawa, S. The aftereffect of the ensemble average of facial expressions on subsequent facial expression recognition. *Attention Percept. Psychophys.* **84**(3), 815–828. <https://doi.org/10.3758/s13414-021-02407-w> (2022).
63. Burton, N., Jeffery, L., Skinner, A. L., Benton, C. P. & Rhodes, G. Nine-year-old children use norm-based coding to visually represent facial expression. *J. Exp. Psychol. Hum. Percept. Perform.* **39**(5), 1261–1269. <https://doi.org/10.1037/a0031117> (2013).
64. Cook, R., Matei, M. & Johnston, A. Exploring expression space: Adaptation to orthogonal and anti-expressions. *J. Vis.* **11**(4), 2. <https://doi.org/10.1167/11.4.2> (2011).
65. Rhodes, G. *et al.* Aftereffects support opponent coding of expression. *J. Exp. Psychol. Hum. Perception Perform.* **43**(3), 619–628. <https://doi.org/10.1037/xhp0000322> (2017).
66. Skinner, A. L. & Benton, C. P. Anti-expression aftereffects reveal prototype-referenced coding of facial expressions. *Psychol. Sci.* **21**(9), 1248–1253. <https://doi.org/10.1177/0956797610380702> (2010).
67. Keating, C. T., Sowden, S. & Cook, J. L. Comparing internal representations of facial expression kinematics between autistic and non-autistic adults. *Autism Res.* **15**(3), 493–506. <https://doi.org/10.1002/aur.2642> (2022).
68. Keating, C. T., & Cook, J. L. Sadness, sorrow or despair: Improving existing tasks assessing emotional granularity. *PsyPAG Q.* (2023).
69. Wang, Y., Shangguan, C., Gu, C. & Hu, B. Individual differences in negative emotion differentiation predict resting-state spontaneous emotional regulatory processes. *Front. Psychol.* **11**, 576119. <https://doi.org/10.3389/fpsyg.2020.576119> (2020).
70. Nook, E. C., Sasse, S. F., Lambert, H. K., McLaughlin, K. A. & Somerville, L. H. The nonlinear development of emotion differentiation: Granular emotional experience is low in adolescence. *Psychol. Sci.* **29**(8), 1346–1357. <https://doi.org/10.1177/0956797618773357> (2018).
71. Barrett, L. F. Feelings or words? Understanding the content in self-report ratings of experienced emotion. *J. Pers. Soc. Psychol.* **87**(2), 266. <https://doi.org/10.1037/0022-3514.87.2.266> (2004).
72. Pond, R. S. Jr. *et al.* Emotion differentiation moderates aggressive tendencies in angry people: A daily diary analysis. *Emotion* **12**(2), 326–337. <https://doi.org/10.1037/a0025762> (2012).
73. Smidt, K. E. & Suvak, M. K. A brief, but nuanced, review of emotional granularity and emotion differentiation research. *Curr. Opin. Psychol.* **3**, 48–51. <https://doi.org/10.1016/j.copsyc.2015.02.007> (2015).
74. Kring, A. M. & Gordon, A. H. Sex differences in emotion: Expression, experience, and physiology. *J. Pers. Soc. Psychol.* **74**(3), 686–703. <https://doi.org/10.1037/0022-3514.74.3.686> (1998).
75. Schirmer, A. Sex differences in emotion. *The Cambridge handbook of human affective neuroscience*, 591–610 (2013).
76. Thompson, A. E. & Voyer, D. Sex differences in the ability to recognise non-verbal displays of emotion: A meta-analysis. *Cogn. Emotion* **28**(7), 1164–1195. <https://doi.org/10.1080/02699931.2013.875889> (2014).
77. Davis, E. *et al.* Emotion experience and regulation in China and the United States: how do culture and gender shape emotion responding?. *International Journal of Psychology* **47**(3), 230–239. <https://doi.org/10.1080/00207594.2011.626043> (2012).
78. Mesquita, B. & Karasawa, M. Different emotional lives. *Cogn. Emotion* **16**(1), 127–141. <https://doi.org/10.1080/0269993014000176> (2002).
79. Mesquita, B. & Walker, R. Cultural differences in emotions: A context for interpreting emotional experiences. *Behav. Res. Ther.* **41**(7), 777–793. [https://doi.org/10.1016/S0005-7967\(02\)00189-4](https://doi.org/10.1016/S0005-7967(02)00189-4) (2003).
80. Mesquita, B., Frijda, N. H. & Scherer, K. R. Culture and emotion. *Handb. Cross Cult. Psychol. Basic Process. Hum. Dev.* **2**, 255 (1997).
81. Mesquita, B. Emotions as dynamic cultural phenomena. In R. Davidson, H. Goldsmith, & P. Rozin (Eds.), *The handbook of the affective sciences* (Oxford University Press, New York, 2003).
82. Jack, R. E., Caldara, R. & Schyns, P. G. Internal representations reveal cultural diversity in expectations of facial expressions of emotion. *J. Exp. Psychol. Gen.* **141**(1), 19. <https://doi.org/10.1037/a0023463> (2012).
83. Jack, R. E., Sun, W., Delis, I., Garrod, O. G. & Schyns, P. G. Four not six: Revealing culturally common facial expressions of emotion. *J. Exp. Psychol. Gen.* **145**(6), 708 (2016).
84. Jack, R. E., Garrod, O. G., Yu, H., Caldara, R. & Schyns, P. G. Facial expressions of emotion are not culturally universal. *Proc. Natl. Acad. Sci.* **109**(19), 7241–7244. <https://doi.org/10.1073/pnas.1200155109> (2012).
85. Chen, C. *et al.* Distinct facial expressions represent pain and pleasure across cultures. *Proc. Natl. Acad. Sci.* **115**(43), E10013–E10021. <https://doi.org/10.1073/pnas.1807862115> (2018).
86. Faul, F., Erdfelder, E., Lang, A. G. & Buchner, A. G\* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* **39**(2), 175–191. <https://doi.org/10.3758/BF03193146> (2007).
87. Button, K. S. *et al.* Power failure: Why small sample size undermines the reliability of neuroscience. *Nat. Rev. Neurosci.* **14**(5), 365–376. <https://doi.org/10.1038/nrn3475> (2013).
88. Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J. & Clubley, E. The autism-spectrum quotient (AQ): Evidence from asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *J. Autism Dev. Disord.* **31**(1), 5–17. <https://doi.org/10.1023/A:1005653411471> (2001).
89. Bagby, R. M., Parker, J. D. A. & Taylor, G. J. The twenty-item Toronto alexithymia scale-I Item selection and cross-validation of the factor structure. *J. Psychos. Res.* **38**(1), 23–32. [https://doi.org/10.1016/0022-3999\(94\)90005-1](https://doi.org/10.1016/0022-3999(94)90005-1) (1994).
90. Chierchia, G. *et al.* The matrix reasoning item bank (MaRS-IB): Novel, open-access abstract reasoning items for adolescents and adults. *R. Soc. Open Sci.* **6**(10), 190232. <https://doi.org/10.1098/rsos.190232> (2019).



91. Keating, C. T. & Cook, J. An emotional granularity task without labels: Abstracts from the 2022 Annual Meeting of the Society for Affective Science. *Health. Affect. Sci* **4**, 210–222. <https://doi.org/10.1007/s42761-022-00174-x> (2023).
92. van der Klis, M. & Tellings, J. Generating semantic maps through multidimensional scaling: linguistic applications and theory. *Corpus Ling. Ling. Theory* <https://doi.org/10.1515/clt-2021-0018> (2022).
93. Huggins, C. F., Cameron, I. M. & Williams, J. H. Autistic traits predict underestimation of emotional abilities. *J. Exp. Psychol. Gen.* **150**(5), 930. <https://doi.org/10.1037/xge0000784> (2021).
94. Kreidler, S. M. *et al.* GLIMMPSE: Online power computation for linear models with and without a baseline covariate. *J. Stat. Softw.* **54**(10), 1–26. <https://doi.org/10.18637/jss.v054.i10> (2013).
95. Asendorpf, J. B., Conner, M., de Fruyt, F., De Houwer, J., Denissen, J. J. A., Fiedler, K., Fiedler, S., Funder, D. C., Kliegl, R., Nosek, B. A., Perugini, M., Roberts, B. W., Schmitt, M., van Aken, M. A. G., Weber, H., & Wicherts, J. M. Recommendations for increasing replicability in psychology. In A. E. Kazdin (Ed.), *Methodological issues and strategies in clinical research* (pp. 607–622). American Psychological Association. <https://doi.org/10.1037/14805-038> (2016).
96. Jack, R. E., Garrod, O. G. & Schyns, P. G. Dynamic facial expressions of emotion transmit an evolving hierarchy of signals over time. *Curr. Biol.* **24**(2), 187–192. <https://doi.org/10.1016/j.cub.2013.11.064> (2014).
97. Keating, C. T. & Cook, J. L. ExpressionMap: A novel method for indexing features of visual emotion representations. *Cogn. Psychol. Bull.* **8**, 1 (2023).
98. Keating, C. T. & Cook, J. L. It's all in the mind: linking internal representations of emotion with facial expression recognition. *Cogn. Psychol. Bull.* **7**, 61–63 (2022).
99. Kenward, M. G. & Roger, J. H. Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics* **1**, 983–997 (1997).
100. Luke, S. G. Evaluating significance in linear mixed-effects models in R. *Behav. Res. Methods* **49**, 1494–1502. <https://doi.org/10.3758/s13428-016-0809-y> (2017).
101. Lee, M. D. & Wagenmakers, E.-J. Bayesian cognitive modelling: A practical course. *Cambridge University Press* <https://doi.org/10.1017/CBO9781139087759> (2014).

### Author contributions

C.T.K. and J.L.C. conceptualised and designed the study. C.T.K. collected the data. C.T.K. analysed the data, wrote an initial draft, and produced the figures. C.T.K. and J.L.C. reviewed and edited the initial draft. Supervision was conducted by J.L.C. Both authors reviewed and provided feedback on the draft and approved the submitted manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-48469-8>.

**Correspondence** and requests for materials should be addressed to C.T.K.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023