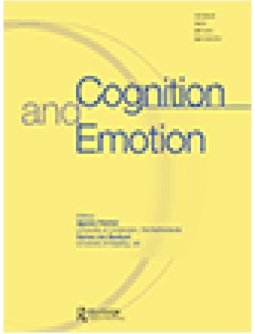


Cognition and Emotion

ISSN: 0269-9931 (Print) 1464-0600 (Online) Journal homepage: http://www.tandfonline.com/loi/pcem20

Emotion perception from a componential perspective

Vera Shuman, Elizabeth Clark-Polner, Ben Meuleman, David Sander & Klaus R. Scherer

To cite this article: Vera Shuman, Elizabeth Clark-Polner, Ben Meuleman, David Sander & Klaus R. Scherer (2015): Emotion perception from a componential perspective, Cognition and Emotion, DOI: 10.1080/02699931.2015.1075964

To link to this article: http://dx.doi.org/10.1080/02699931.2015.1075964

Published online: 26 Aug 2015.

COGNITION AND EMOTION, 2015
http://dx.doi.org/10.1080/02699931.2015.1075964

Emotion perception from a componential perspective

Vera Shuman[a†], Elizabeth Clark-Polner[a,b†], Ben Meuleman[a], David Sander[a,b] and Klaus R. Scherer[a]

[a]Swiss Center for Affective Sciences, University of Geneva, Geneva, Switzerland; [b]Department of Psychology, University of Geneva, Geneva, Switzerland

ABSTRACT
The common conceptual understanding of emotion is that they are multi-componential, including subjective feelings, appraisals, psychophysiological activation, action tendencies, and motor expressions. Emotion perception, however, has traditionally been studied in terms of emotion labels, such as "happy", which do not clearly indicate whether one, some, or all emotion components are perceived. We examine whether emotion percepts are multi-componential and extend previous research by using more ecologically valid, dynamic, and multimodal stimuli and an alternative response measure. The results demonstrate that observers can reliably infer multiple types of information (subjective feelings, appraisals, action tendencies, and social messages) from complex emotion expressions. Furthermore, this finding appears to be robust to changes in response items. The results are discussed in light of their implications for research on emotion perception.

ARTICLE HISTORY
Received 31 October 2014
Revised 17 July 2015
Accepted 20 July 2015

KEYWORDS
Emotion perception; componential emotion; emotion expression

Research on emotion perception has long used the same basic template: participants are presented with stimuli depicting different expressions and asked to identify the emotion by indicating the label with which the expressions are associated. "Emotions", in these paradigms, are operationalised with a list of colloquial, categorical, descriptors (e.g., "happy" and "sad"), from which participants can select the most appropriate label(s). This is the standard, not only in studies of facial, vocal, and bodily expressions, but also in cross-cultural, developmental, and clinical emotion research (Aviezer et al., 2009; Baum & Nowicki, 1998; Coulson, 2004; Fiorentini, Schmidt, & Viviani, 2012; Naab & Russell, 2007; as discussed in Scherer, Clark-Polner, & Mortillaro, 2011).

These emotion labels, however, lack clarity with regard to whether one, some, or all emotion components are perceived. Emotions are commonly regarded as multi-componential, including cognitive appraisals of the situation, subjective feelings, and action tendencies (Scherer & Grandjean, 2008). Additionally, expressions can function as social messages to communicate information to perceivers (Fridlund, 1992; Yik & Russell, 1999). Labels such as "happiness" may relate to emotion components in various ways.

For example, emotion labels may be more strongly associated with subjective feelings than with other components. If this were true, then using emotion labels would lead to inappropriate inferences about other emotion components. This issue is important because emotion perception research is often used as the basis for claims about the nature of emotion itself. The emotion perception literature has been described, for example, as providing "impressive evidence" (Izard, 1980) for the discrete theories of emotion. Such inferences have also been called "definitive" (Ekman, 1980), "conclusive" (Oster, Daily, & Goldenthal, 1989), and "sufficiently robust [so as to be considered] … an established axiom of behavioral science" (Izard & Saxton, 1988). Based on the current literature alone, however, we cannot conclude that these "facts" will hold up if we reconcile our experimental operationalisation of emotion percepts with our

CONTACT Vera Shuman vsachari@umich.edu
[†]Vera Shuman is now at Saint Paul, MN, and Elizabeth Clark-Polner is now at the Department of Psychology, University of Chicago. Both authors contributed equally.

conceptual understanding about the type of information potentially conveyed by emotion expressions.

Alternatively to being predominantly associated with subjective feelings, emotion labels may subsume the entire emotion concept and be associated with any of the emotion components. There still remains a problem with using general emotion labels. If emotion labels indicate the presence of all components, assessing emotion percepts with these labels may obscure which component is most relevant in a given context. For example, in negotiations, the appraisals, action tendencies, and social messages inferred from emotion expressions may have more impact than conscious feelings on the dynamic between negotiators. Such specific inferences cannot be appropriately studied by using general emotion labels, but only by using component-specific response items.

Prior research

Three prior empirical studies have been aimed at addressing the question of what type of information is perceived when viewing emotion expressions (Horstmann, 2003; Scherer & Grandjean, 2008; Yik & Russell, 1999). In each case, researchers asked participants to label emotion expressions using items reflecting different types of emotion-relevant information (subjective feelings, social messages, and action tendencies). Scherer and Grandjean (2008) additionally included appraisal items. The authors in these studies computed hit rates: the proportions of responses that matched experimenters' a priori predictions. In each case, the authors found that all emotion components achieved above chance hit rates, and, in Horstmann's and in Scherer and Grandjean's study, that subjective feeling labels achieved the highest hit rates (e.g., Scherer & Grandjean's findings are displayed in Figure 1(a)). This suggests that perceivers may indeed reliably interpret emotion expressions as conveying multiple types of emotion-relevant information.

Several methodological aspects of the previous studies need to be improved, however, before strong conclusions can be drawn: first, neither Yik and Russell (1999) nor Horstmann (2003) include a "none of the above" option in their forced-choice response measure, an omission that has been demonstrated to inflate hit rates (Russell, 1994).

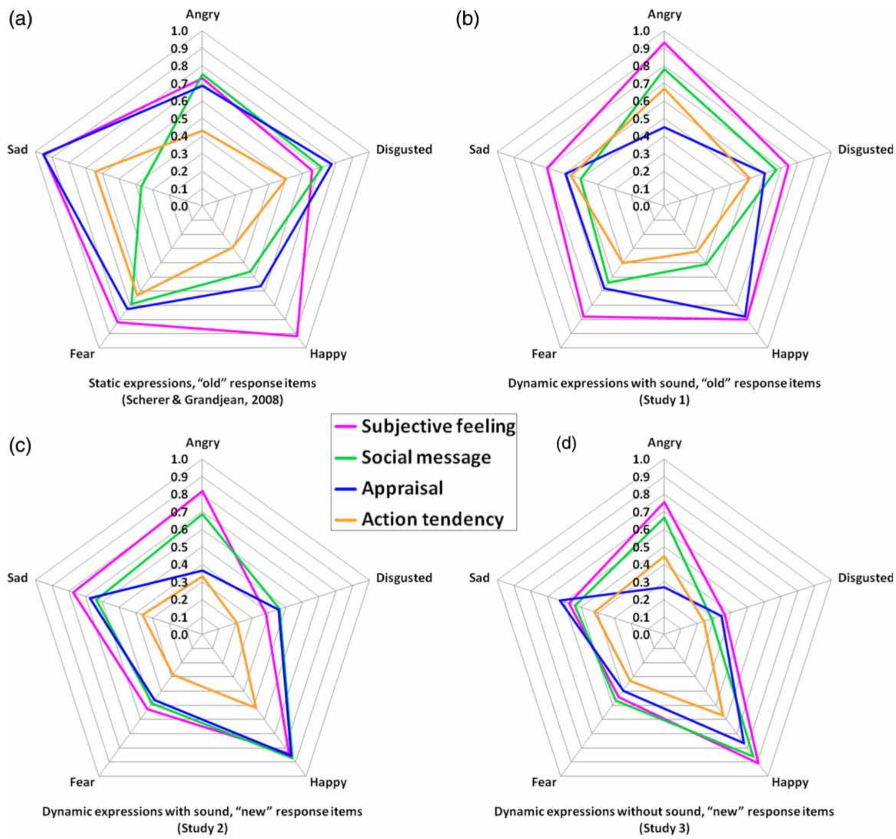
Second, each used as stimuli still photographs from Ekman and Friesen's *Pictures of Facial Affect* (1976). This is important because the perceptual characteristics that differentiate still pictures from the types of expressions we see in real life—dynamic displays with sound—appear to play an important role in the recognition and evaluation of emotion expressions. Whether hit rates for dynamic stimuli are higher than for static stimuli depends on various factors; with natural faces, there appears to be no systematic advantage of dynamic expressions (Krumhuber, Kappas, & Manstead, 2013). However, research on brain-damaged and autistic individuals as well as neuroimaging studies suggests that dynamic emotion expressions may be processed via different perceptual mechanisms compared to static stimuli (Back, Ropar, & Mitchell, 2007; Harwood, Hall, & Shinkfield, 1999; Kilts, Egan, Gideon, Ely, & Hoffman, 2003; Sato, Kochiyama, Yoshikawa, Naito, & Matsumura, 2004; Trautmann, Fehr, & Herrmann, 2009). The emotion perception systems engaged in the studies by Horstmann (2003), Scherer and Grandjean (2008), and Yik and Russell (1999) might thus be different than those we rely on in everyday life.

Finally, the response items used in previous studies (Scherer & Grandjean, 2008; Yik & Russell, 1999) confound the *type of information* expressed by the labels in the different conditions and the *specificity* of those labels. For example, the subjective feeling item for sadness is simply "sad", whereas the appraisal item is "I just lost someone very close to me" (Figure 1). The subjective feeling item appears to refer to all instances of sadness; the appraisal item, in contrast, indexes a very specific instance of sadness—that elicited by the loss of a close other. Any previously reported similarities or differences in hit rates may thus be difficult to interpret.

Present studies

The objective of the research presented here was to address the issue of how to best measure inferences from emotion expressions, while trying to improve on the methodology used in some of the prior research. Specifically, we examined the following two research questions in three studies.

First, are recognition hit rates for different emotion components still above chance if expression stimuli are more perceptually complex than in previous research? We examined dynamic visual stimuli with audio content (Studies 1 and 2) and without audio content (Study 3). Our hypothesis was that hit rates



Response items	Subjective Feeling	Social message	Action tendency	Appraisal
"Old"	Sadness	Please hold me and comfort me	I feel like weeping	I just lost someone very close to me
	Happiness	Hi, how are you doing?	I want to jump up and down	I have just received a gift
	Fear	Don't hurt me! I give up	I feel like getting out	I am in a dangerous situation and I don't know how to get out of it
	Disgust	That stinks!	I try to avoid smelling this stuff	This is really unpleasant
	Anger	Back off or I'll attack	I want to hit this person	This blocks my plans but I will fight to get what I want
"New"	I feel sad	You make me sad	I might cry	That is sad
	I feel happy	You make me happy	I might approach	That is great
	I feel scared	You scare me	I might run	That is dangerous
	I feel disgusted	You disgust me	I might withdraw	That is disgusting
	I feel angry	You make me angry	I might lash out	That is unacceptable

Figure 1. Above: Model based estimates of mean hit rates in % by emotion expression and component condition. Values on the vertical axis of each graph represent the proportion of trials for which the target response was given. Below: Response items used in Studies 1–3. "Old" items are those items used previously by Scherer and Grandjean (2008) and Yik and Russell (1999) and used in study 1; table adapted from Scherer and Grandjean (2008; p. 790). Note that in Yik & Russell (1999), only subjective feeling and social message items were used. "New" items indicate the revised response items used in studies 2 and 3.

would still be above chance when stimuli were more perceptually complex than in previous research.

Second, are hit rates still significantly higher in the subjective feeling condition than in other conditions when the response measure is adapted such that the labels across conditions differ less in specificity than in previous research? We developed a new set of items that were more similar in specificity as well as in length using grammatical features. Studies 2 and 3 tested hit rates with these new items. In all studies, we included a "none of the above" response option.

Methods

Participants

Participants were students and community members participating for partial fulfilment of course requirements or monetary compensation. Sixty-one individuals (48 women) took part in Study 1 (*M* age = 37.4); 112 (53 women) took part in Study 2; and 158 (76 women) took part in Study 3. Sample sizes were determined based on past research using this paradigm (Scherer & Grandjean, 2008; Yik & Russell, 1999).

Design

In Study 1, each participant was randomly assigned to two of the four emotion component conditions. The experiment thus utilised a 4 (component) × 5 (expression) design, with emotion component (subjective feeling, appraisal, social message, or action tendency) as a partial within-subjects variable, and expression (happiness, sadness, fear, anger, and disgust) as a full within-subjects variable. In Studies 2 and 3, participants were assigned to one of the four component conditions (emotion component was thus a between-subjects variable).

Procedure

In each study, participants first filled out a demographics questionnaire and completed a practice session. They were then shown 60 stimuli in random order. For each trial, participants were asked to identify what they saw in the clip by choosing a label from an answer bank. The response items reflected feelings, appraisals, social messages, or action tendencies associated with each of the five expression categories. Each answer bank also included a "none of the above" option. Dependent variables were participants' hit rates and response times. Participants completed all sessions online.

Stimuli

Sixty dynamic audio/video clips of emotion expressions were selected from the Geneva Multimodal Emotion Portrayals corpus (GEMEP; Bänziger & Scherer, 2010, see Figure 2). This corpus contains 1260 audio-visual recordings of expressions of 18 different emotions encoded by 10 professional theatre actors. Twelve portrayals of each of the five expression categories were used in the study (anger, disgust, fear, happiness, sadness; chosen based on those included in previous research). The portrayals selected were balanced by actor gender and audio content (two pseudo-linguistic sentences and one sustained vowel, "aaa").

This stimulus set has several advantages. First, it offers emotion expression stimuli that are more realistic—in terms of their perceptual characteristics—than those used previously, while still allowing for a level of standardisation that is unattainable when using "natural" (spontaneously occurring) emotion expressions. Unlike "natural" stimuli, however, the GEMEP stimuli include vocal expression without verbal content, which is known to bias perceptions (Gendron, Lindquist, Barsalou, & Barrett, 2012).

The portrayals used here were taken from a subset of stimuli (the "core set"). The primary criterion in the selection of the core set was a high believability rating. A secondary criterion was clarity, defined by hit rates in an emotion perception task in which raters labelled the portrayals using traditional emotion labels. This clarity criterion helped to ensure that the portrayals were more similar within than between groups, but, as this was not the sole selection criterion, the filter it applied was not too strong; indeed, validation studies for the stimuli within the core set report hit rates with means and variances similar to studies using real, *in vivo*, expressions (Bänziger, Mortillaro, & Scherer, 2012; Bänziger & Scherer, 2010).

Response measures

In Study 1, participants were asked to choose the best label for the stimuli they saw from the items developed by Yik and Russell (1999) and Scherer and Grandjean (2008). In Studies 2 and 3, choice options were revised response items (Figure 1).

Figure 2. Still frames from GEMEP corpus dynamic video stimuli. Stimuli were chosen randomly from the GEMEP "core set", a subset of stimuli selected for high believability and authenticity. A detailed description of the production and validation of the GEMEP corpus can be found in Bänziger et al. (2012), and at http://www.affective-sciences.org/gemep. Example of the multimodal stimuli—with different response options than the ones used in the current study—can be seen in this demo (http://www.unige.ch/cisa/properemo/gert/demo.php).

Our primary goal when revising the response items was to reduce differences in specificity and length between components and between emotions. This was done to reduce the potential confounding of our manipulation of interest (emotion components). We made all response options more general, as opposed to more specific. This was done to reduce the potential "mismatch" between the specific instance of an emotion described by a response option and the instance(s) of the emotion as envisioned by the actors during stimulus recording.

The primary difference between items for different components is now grammatical: items in the subjective feeling condition describe how the emotion state feels to the emoter in the first person (e.g., "I feel scared"; "I feel disgusted"). In the social message condition, the focus of the sentence is the relation between observer and expresser, and the observer is the subject of the sentence ("You scare me"; "You disgust me"). Note that this social message condition is less closely aligned with Fridlund's (1992) social motive concept, which includes action readiness, action requests, and other messages. The old social message items were more closely aligned with Fridlund's social motive concept, but they were less clearly differentiated from the action tendency condition. Some of the meaning of Fridlund's social motive concept is captured with the new action tendency items. In the action tendency condition, the emoter is again the subject of the sentence, but instead of an adjective describing a feeling at the end, there is a verb phrase describing an action tendency (e.g., "I might run"; "I might withdraw"). Finally, in the appraisal condition, the subject of the sentence is the pronoun "that", which indicates an external object or situation ("That is dangerous"; "That is disgusting"). Items within a component condition thus share the same construction and differ only in the adjective or adverb used (with adjective and adverbs designed to be as general as possible). Items within an emotion condition differ primarily in their grammatical construction, and there is less variation in specificity.

To ensure that participants were interpreting these items in the same way as the authors, we conducted a survey study in which we asked participants to generate phrases to describe each emotion for each component (e.g., "How might you convey disgust in a *social situation*?"; "How might you convey disgust in terms of one's *actions*?"). Each of the items used in our revised item set appeared at least once on these lists in the same or in a substantially similar form (e.g., "I might withdraw"; "I might back away").

Results

Data analytic strategy

Prior to the analysis, irregular observations—defined as trials with a reaction time larger than 20 seconds and trials with a reaction time equal to 0—were removed from the data (3.7%). Data were then averaged across trials for each condition.

Due to the mixed design, traditional analysis methods—for example, repeated measures multivariate analysis of variance—could not be applied in Study 1. Instead, we used a marginal linear mixed model (Fitzmaurice, Laird, & Ware, 2004), which is able to account for missing data and flexibly model the structure of the within-subjects covariance matrix. Supplementary materials contain further information on underlying assumptions and the choice of covariance structure (S1). Once a suitable covariance structure was chosen (for these data, a first-order auto-regressive moving average model, see supplementary materials), analyses were conducted according to the procedure for multi-level models outlined by Maxwell and Delaney (2004). For Studies 2 and 3 where the label condition was fully between subjects, repeated measures one-way analyses of variance were used to analyse the data. Supplementary materials contain data analysis results concerning reaction times (S2).

We also compared hit rates to chance levels. In order to account for response bias, we compared hit rates to the empirical chance levels rather than the theoretical chance levels using R statistical software (version 2.14.1). The empirical chance level of a target response for a given expression was estimated by calculating the overall proportion of choices for that expression (regardless of whether it was correct or incorrect) within each emotion component. We then applied a non-parametric one-sample permutation test to the (trial-averaged) data, for each condition testing whether the mean hit rate was significantly different from the empirical chance level.

Finally, in Study 1, we calculated unbiased hit rates. This allowed us to assess the degree to which any significant effects found here may be an artefact, attributable to participants tending to choose specific emotions in different conditions, regardless of what is actually expressed.

Study 1

Hit rates for traditional items in all conditions were well above chance, with all permutation *p*-values exactly equal to 0. Model-based estimates of mean hit rates are displayed in Figure 1(b) and ranged from 33% to 93% with an average of 62% (Table S1 in supplementary materials). Average hit rates ranged for emotion components from 51% (action tendencies) to 79% (subjective feelings), and for emotion expressions from 57% (fear) to 71% (anger). “None of the above” rates ranged from 1.7% (anger, subjective feeling) to 61% (happiness, action tendency) with an average of 23% (see supplementary materials, Table S2).

We found a significant interaction between emotion component and expression, $F(12, 228) = 8.47$, $p < 0.001$. Simple effects tests revealed differences in hit rates between emotion components within each expression except sadness, with $F(3, 228) = 19.58$, $p_{adj} < 0.001$, for anger, $F(3, 228) = 4.42$, $p_{adj} = 0.024$, for disgust, $F(3, 228) = 11.67$, $p_{adj} < 0.001$, for fear, $F(3, 228) = 29.58$, $p_{adj} < 0.001$, for happiness, and $F(3, 228) = 3.52$, $p_{adj} = 0.08$, for sadness, respectively. The results of all post hoc tests of differences within each expression category between components showed 15 significant contrasts with varied differences across components (see supplemental materials, Tables S3a and S4a). High hit rates for subjective feelings were implicated in 9 out of 15 significant contrasts, and low hit rates for action tendencies in 7 contrasts. Anger was additionally associated with low hit rates for appraisals.

Study 2

Model-based estimates of mean hit rates for the revised response items are displayed in Figure 1(c) and ranged from 21% to 87% with an average of 55% (Table S1 in supplementary materials). Average hit rates ranged for emotion components from 34% (action tendencies) to 67% (subjective feelings), and for emotion expressions from 38% (disgust) to 77% (happiness). “None of the above” rates ranged from 2% (anger, subjective feeling) to 17% (happiness, action tendency) with an average of 8%.

We found a significant interaction between emotion expression and component, $F(12, 400) = 4.10$, $p < 0.001$. Subsequent simple effects tests revealed significant effects of emotion component within each of the five types of expression conditions, with $F(3, 106) = 15.89$, $p_{adj} < 0.001$, for anger, $F(3, 102) = 6.70$, $p_{adj} < 0.001$, for disgust, $F(3, 105) = 9.86$, $p_{adj} < 0.001$, for happiness, $F(3, 103) = 4.71$, $p_{adj} \leq 0.05$, for fear, and $F(3, 106) = 15.89$, $p_{adj} < 0.001$, for sadness. As in Study 1, hit rates for subjective feeling labels were high; however, they differed significantly from rates for other components in only 5 out of 14 contrasts, substantially fewer than in Study 1. Similar to Study 1, hit rates for action tendencies were significantly lower than for other components in 12 contrasts. Anger was again associated with low hit rates for appraisals. Post hoc tests for contrasts between specific components within each expression category are reported in supplemental materials (Tables S3b and S4b).

Study 3

Model-based estimates of mean hit rates for stimuli without sound ranged from 24% to 87% with an average of 51% (Figure 1(d) and Table S1). Hit rates for emotion components ranged from 40% (action tendencies) to 61% (subjective feelings), and for emotion expressions from 31% (disgust) to 78% (happiness).

As in Studies 1 and 2, the interaction between emotion expression and component was significant, $F(12, 556) = 6.552$, $p < 0.001$. Subsequent simple effects tests revealed significant effects of emotion component for anger, $F(3, 149) = 23.47$, $p_{adj} < 0.001$, and happiness, $F(3, 146) = 8.35$, $p_{adj} < 0.001$, but not for disgust, $F(3, 146) = 0.94$, $p_{adj} = 1$, fear, $F(3, 145) = 1.39$, $p_{adj} = 1$, or sadness, $F(3, 147) = 3.27$, $p_{adj} = 0.115$. The number of significant contrasts was seven, much lower than in Studies 1 and 2. Subjective feelings were involved in three contrasts. Similar to Studies 1 and 2, the effects for anger and happiness appeared to be driven by low hit rates for action tendencies in five significant contrasts; for anger, differences were again due to lower hit rates for appraisals, comprising the remaining two significant contrasts. Post hoc tests for contrasts between specific components within each expression condition are reported in supplemental materials (Tables S3c and S4c).

Unbiased hit rates

A sample confusion matrix was constructed containing data from Study 1 (Table S2 in supplementary materials). The matrices indicate that respondents were biased in their choice of response by condition. In the social message condition, for example, participants tended to choose the response options associated with anger and disgust more often than the option associated with happiness. To account for these biases, we also calculated unbiased hit rates for each expression in each component, using the procedure outlined by Russell and Fernandez-Dols (1997). Bias correction did not change the pattern of the results. See the supplementary materials (Table S5).

Discussion and conclusion

Previous research has suggested that emotion expressions can convey multiple types of information, including subjective feelings, social messages, action tendencies, and appraisals. However, our ability to draw conclusions has been limited by both the stimuli and response measures used.

Our first research question was whether recognition hit rates would still be above chance if expressions were enacted and dynamic rather than posed and static. The results from all three studies confirm the prediction that participants would be able to perceive information about each of the theorised components of emotion, including subjective experience, action tendencies, social messages, and appraisals, with hit rates greater than those which would be expected by chance. The data reported here demonstrate that the finding that emotion percepts contain information relevant to multiple emotion components is robust to variation in the perceptual characteristics of the emotion expressions being considered, and additionally to variation in the nature of the response items used to measure expression recognition.

The second research question was whether hit rates would still be significantly higher in the subjective feeling condition than in other conditions when the response measure was adapted such that the labels across conditions differed less in specificity than in previous research. With more general response items, the number of significant comparisons with higher hit rates for subjective feeling items dropped from nine in Study 1 to five and three in Studies 2 and 3, respectively, suggesting that other components can be as strongly associated with emotion expressions as subjective feelings. At the same time, no emotion component exceeded subjective feelings in hit rate. In other words, subjective feelings were always among the emotion components most strongly associated with the emotion expressions.

Several additional observations raise more questions than answers. First, in all studies, action tendencies were the least well-recognised component. In Studies 2 and 3, action tendency items were formulated in such a way as to correspond to the items for other components. The grammatical structure was the same as with all other items, and items were similar to appraisal items for happiness, fear, and anger in that they did not contain the emotion word (e.g., "that is great" *vs*. "that is *sad*"). The items ranged from concrete behaviours ("I might cry") to general tendencies ("I might approach"). We cannot rule out that the results are due to the items used in the studies. Alternatively, facial expressions lacking situational context may indeed not convey action tendencies well compared to other components.

Second, among the expressions, anger stood out with greater differences in hit rates across conditions: subjective feelings and social messages were both better recognised than appraisals (Studies 1–3) and action tendencies (Studies 2 and 3). The result can be attributed to confusions with other items rather than increased "none of the above" responses. For example, in Study 1, the response rate for appraisals was 30% for the disgust item. Anger and disgust are often confused in facial expression research. A simple interpretation is not warranted, however, because confusion rates between anger and disgust were low for subjective feeling and social message items, and because confusions were different in the other studies (e.g., in Study 2, anger appraisals were confused with fear appraisals in 34% of responses). Interpretations are further impeded because Scherer and Grandjean (2008), who used the same response items as in Study 1 but different stimuli, found more confusions of the anger with the sadness items for appraisals and action tendencies. The current findings suggest that confusion rates are influenced by expression characteristics (comparison of Study 1 and previous research) and response characteristics (comparison of Studies 1 and 2). Further research is needed to analyse in detail which features of stimulus and response options drive confusions between emotions.

Third, we observed variable endorsement of the "none of the above" item, with higher endorsement in Study 1 than in Study 2. Comparisons of Study 1 with Scherer and Grandjean's (2008) study showed overall higher response rates for "none of the above" in Study 1. Together, this suggests that the level of endorsement of this response option varies with the specific response items (comparison of Study 1 and Study 2) and with the stimulus material (comparison of Study 1 and previous research). At the same time, the rank order of "none of the above" endorsement for emotion component conditions was similar across studies, increasing from subjective feeling and appraisal to action tendency and social message labels. Similar to Scherer and Grandjean, "none of the above" was endorsed most often for happiness expressions in Study 1 in the social message and action tendency conditions, mirroring low hit rates in these conditions. In contrast, in previous studies without "none of the above" response options, hit rates were much higher (Yik & Russell, 1999). This underscores the importance of including a "none of the above" response option.

Fourth, in Study 3 where vocal expressions were absent we found the smallest number of differences in hit rates between emotion components; future research is needed to examine whether the observed difference between studies would be significant in a direct comparison of voice/no voice conditions.

Finally, in Studies 2 and 3 where we added the stem "I feel" to the canonical emotion labels, we obtained lower hit rates than in Study 1, where we used the emotion labels alone. This may suggest that the canonical emotion terms capture more than the subjective feeling component of emotions, and supports a multi-componential understanding of the emotion concept.

Given the limitations of any single set of response items, we cannot infer that the specific patterns reported above will generalise beyond the items used here. The items in Studies 2 and 3 were closer in specificity than in previous research, but not fully equivalent. As pointed out by a reviewer, "I feel sad" subsumes situations where "you make me feel sad", "someone else makes me feel sad", and "something makes me feel sad". In our studies, we combined grammatical criteria and analysis of participant-generated responses to generate items with similar specificity, length, and construct validity. One could achieve even greater similarity in specificity, though at the expense of similarity in length and readability (e.g., "you/something/someone makes me feel sad"). Future research may benefit from alternative methods of item generation and validation.

Another limitation of our studies is that it is impossible to say whether these results will generalise to spontaneous expressions. In the current studies using the GEMEP dynamic multimodal stimuli, the expressions were obtained from professional actors who were instructed to enter into different emotion states on command. In using these stimuli, though we achieved more ecological validity than in previous studies using static posed expressions, we did not achieve the highest degree of ecological validity possible. Researchers who wish to further increase ecological validity by studying spontaneous expressions loose standardisation across stimuli. Our results may complement such research by focusing attention to patterns that replicate across study approaches, and the current results may thereby facilitate the interpretation of future research findings from studies with little experimental control.

As mentioned above, a question for future research is a direct comparison of hit rates across stimulus features. Future research should therefore manipulate static *vs.* dynamic stimuli within one study. Additionally, the underlying process of emotion recognition could be further examined by studies systematically manipulating dynamic stimuli to examine at what feature constellation of a dynamic display individuals select particular responses (cf. Fiorentini et al., 2012).

The question of the way in which we recognise emotion expressions and the types of inferences we make when we view them is relevant for research on emotion perception such as emotion processing in the brain. For example, imaging findings are more often than not interpreted narrowly in terms of how the structure or network identified may be involved in our deduction of how the target person feels. However, unless otherwise instructed, a respondent may consider instead what might have elicited the person's emotion, how he or she is likely to behave, or what he or she is trying to communicate to observers, all of which may activate different brain regions resulting in erroneous conclusions.

To conclude, the results support the notion that emotion percepts may represent several types of information relevant to emotion states. Additional research is needed to further elucidate the relationship between these different types of inferences and the emotion expressions on which they are based.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was supported by an ERC Advanced Grant in the European Community's 7th Framework Program under grant agreement 230331-PROPEREMO (Production and perception of emotion: an affective sciences approach) to Klaus Scherer and by the National Center of Competence in Research (NCCR) Affective Sciences financed by the Swiss National Science Foundation (51NF40-104897) and hosted by the University of Geneva.

Supplemental data

Supplemental data for this article can be accessed at doi:10.1080/02699931.2015.1075964.

References

Aviezer, H., Bentin, S., Hassin, R. R., Meschino, W. S., Kennedy, J., Grewal, S., … Moscovitch, M. (2009). Not on the face alone: Perception of contextualized face expressions in Huntington's disease. *Brain: A Journal of Neurology*, *132*, 1633–1644.

Back, E., Ropar, D., & Mitchell, P. (2007). Do the eyes have it? Inferring mental states from animated faces in autism. *Child Development*, *78*, 397–411.

Bänziger, T., Mortillaro, M., & Scherer, K. R. (2012). Introducing the Geneva Multimodal Expression Corpus for experimental research on emotion expression. *Emotion*, *12*, 1161–1179.

Bänziger, T., & Scherer, K. R. (2010). Introducing the Geneva Multimodal Emotion Portrayal (GEMEP) Corpus. In K. R. Scherer, T. Bänziger, & E. B. Roesch (Eds.), *Blueprint for affective computing: A sourcebook* (pp. 271–294). Oxford: Oxford University Press.

Baum, K. M., & Nowicki Jr, S. (1998). Perception of emotion: Measuring decoding accuracy of adult prosodic cues varying in intensity. *Journal of Nonverbal Behavior*, *22*(2), 89–107.

Coulson, M. (2004). Attributing emotion to static body postures: Recognition accuracy, confusions, and viewpoint dependence. *Journal of Nonverbal Behavior*, *28*(2), 117–139.

Ekman, P. (1980). The face of man: Expressions of universal emotions in a New Guinea village. New York, NY: Garland STPM Press.

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

Fiorentini, C., Schmidt, S., & Viviani, P. (2012). The identification of unfolding facial expressions. *Perception*, *41*(5), 532–555.

Fitzmaurice, G. M., Laird, N. M., & Ware, J. H. (2004). *Applied longitudinal data analysis*. Hoboken, NJ: Wiley.

Fridlund, A. (1992). The behavioral ecology and sociality of human faces. In M. Clark (Ed.), *Review of personality & social psychology* (pp. 90–121). London: Sage Publications.

Gendron, M., Lindquist, K. A., Barsalou, L., & Barrett, L. F. (2012). Emotion words shape emotion percepts. *Emotion*, *12*, 314–325.

Harwood, N. K., Hall, L. J., & Shinkfield, A. J. (1999). Recognition of facial emotional expressions from moving and static displays by individuals with mental retardation. *American Journal of Mental Retardation*, *104*, 270–278.

Horstmann, G. (2003). What do facial expressions convey: Feeling states, behavioral intentions, or action requests? *Emotion*, *3*, 150–166.

Izard, C. E. (1980). Cross-cultural perspectives on emotion and emotion communication. In H. Triandis & W. Lonner (Eds.), *Handbook of cross-cultural psychology: Basic processes* (Vol. 3, pp. 185–222). Boston, MA: Allyn & Bacon.

Izard, C. E., & Saxton, P. M. (1988). Emotions. In R. C. Atkinson, R. J. Herrnstein, G. Lindzey, & R. D. Luce (Eds.), *Stevens' handbook of experimental psychology* (2nd ed., Vol. 1, pp. 627–676). New York, NY: Wiley.

Kilts, C. D., Egan, G., Gideon, D. A., Ely, T. D., & Hoffman, J. M. (2003). Dissociable neural pathways are involved in the recognition of emotion in static and dynamic facial expressions. *NeuroImage*, *18*, 156–168.

Krumhuber, E. G., Kappas, A., & Manstead, A. S. (2013). Effects of dynamic aspects of facial expressions: a review. *Emotion Review*, *5*(1), 41–46.

Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Naab, P. J., & Russell, J. A. (2007). Judgments of emotion from spontaneous facial expressions of New Guineans. *Emotion*, *7*(4), 736–744. http://doi.org/10.1037/1528-3542.7.4.736

Oster, H., Daily, L., & Goldenthal, P. (1989). Processing facial affect. In A. W. Young & H. D. Ellis (Eds.), *Handbook of research on face processing* (pp. 107–161). Amsterdam: Elsevier.

Russell, J. A. (1994). Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychological Bulletin*, *118*, 379–391.

Russell, J. A., & Fernandez-Dols, J. M. (1997). *The psychology of facial expression*. Cambridge: Cambridge University Press.

Sato, W., Kochiyama, T., Yoshikawa, S., Naito, E., & Matsumura, M. (2004). Enhanced neural activity in response to dynamic facial expressions of emotion: An fMRI study. *Cognitive Brain Research*, *20*(1), 81–91.

Scherer, K. R., Clark-Polner, E., & Mortillaro, M. (2011). In the eye of the beholder? Universality and cultural specificity in the expression and perception of emotion. *International Journal of Psychology*, *46*, 401–435.

Scherer, K. R., & Grandjean, D. (2008). Facial expressions allow inference of both emotions and their components. *Cognition & Emotion*, *22*, 789–801.

Trautmann, S. A., Fehr, T., & Herrmann, M. (2009). Emotions in motion: Dynamic compared to static facial expressions of disgust and happiness reveal more widespread emotion-specific activations. *Brain Research*, *1284*, 100–115.

Yik, M., & Russell, J. (1999). Interpretation of faces: A cross-cultural study of a prediction from Fridlund's theory. *Cognition & Emotion*, *13*, 93–104.