Volume 25 – Issues 4-6 – April-June 2017
Visual Cognition

Visual Cognition

ISSN: 1350-6285 (Print) 1464-0716 (Online) Journal homepage: http://www.tandfonline.com/loi/pvis20

Looking beyond the static face in emotion recognition: The informative case of interest

Daniel Dukes, Fabrice Clément, Catherine Audrin & Marcello Mortillaro

To cite this article: Daniel Dukes, Fabrice Clément, Catherine Audrin & Marcello Mortillaro (2017) Looking beyond the static face in emotion recognition: The informative case of interest, Visual Cognition, 25:4-6, 575-588, DOI: 10.1080/13506285.2017.1341441

To link to this article: https://doi.org/10.1080/13506285.2017.1341441

Published online: 18 Jul 2017.

VISUAL COGNITION, 2017
VOL. 25, NOS. 4–6, 575–588
https://doi.org/10.1080/13506285.2017.1341441

Routledge
Taylor & Francis Group

Looking beyond the static face in emotion recognition: The informative case of interest

Daniel Dukes[a,b], Fabrice Clément[a,b], Catherine Audrin[b,c] and Marcello Mortillaro[b]

[a]Cognitive Science Centre, University of Neuchâtel, Neuchâtel, Switzerland; [b]Swiss Center for Affective Sciences, University of Geneva, Geneva, Switzerland; [c]Faculty of Psychology and Education Science, University of Geneva, Geneva, Switzerland

ABSTRACT
A new wave of studies on emotion recognition encourages researchers to look beyond the face as the sole source of pertinent information. One study has proposed that, while there is overwhelming evidence that negative emotions may be differentiated in static facial expressions, postural information is needed to differentiate positive emotions such as pride, love and amusement. We review the most influential of these recent studies before providing some supporting evidence in an exploratory study of a fourth positive emotion – interest – that adds to the recent calls to move beyond static facial expressions of emotion when investigating how we determine how others are feeling and how they are likely to act next.

ARTICLE HISTORY
Received 27 November 2016
Accepted 23 March 2017

KEYWORDS
Body emotion recognition; dynamic emotion recognition; positive emotion; facial emotion recognition; interest

Spontaneous behaviour typically unfolds through multiple, nonverbal modalities at the same time. Most, if not all, emotion researchers would probably agree with this statement. Nevertheless, until very recently, most empirical studies of emotion expression focused on a single modality, usually that of facial expression.

However, a new wave of studies is pushing research on nonverbal emotion expression towards a more comprehensive multimodal approach (Aviezer et al., 2008; de Gelder & Vroomen, 2000; Mortillaro, Mehu, & Scherer, 2013; Scherer & Ellgring, 2007; Van den Stock, Righart, & de Gelder, 2007; Watson et al., 2014). The pioneering work of De Gelder and colleagues on the interaction between the body, the face and the voice in the perception of emotion demonstrated the importance of whole body expressions to alter the perception of emotion from faces and voices (Van den Stock et al., 2007) and the tight relationship between audio and visual processing in emotion perception (de Gelder & Vroomen, 2000). More recently, a similar shift has occurred in developmental studies, underlining how important it is to consider multi-modal emotion recognition in the early stages of emotion development. For example, research has found that 6.5-month-old infants are able to correctly match whole-body and vocal expressions of emotion (Zieber, Kangas, Hock, & Bhatt, 2014) while the ability to recognize emotions from whole-body expressions (without the face) has been demonstrated in pre-schoolers (Nelson & Russell, 2011).

Recent influential studies have not only shown that the combination of facial expressions and bodies can drastically change how even prototypical facial expression of emotions are interpreted (Aviezer et al., 2008), but also that a congruency effect is already present in 6-year-old children (Mondloch, 2012; Mondloch, Horner, & Mian, 2013). Furthermore, research has shown that, in cases of extreme emotional intensity, the valence of the situation cannot be reliably determined from facial expressions alone, but the process requires additional contextual information that perhaps relates to body posture (Aviezer, Trope, & Todorov, 2012; Wenzler, Levine, van Dick, Oertel-Knöchel, & Aviezer, 2016). Cross-modal perceptual influences have also been uncovered in a recent study that found that emotional information in an unattended modality can disambiguate the perceived emotion of a crowd (McHugh, Kearney, Rice, & Newell, 2012). Other studies have found cross-modal embodiment effects between

CONTACT Marcello Mortillaro marcello.mortillaro@unige.ch Swiss Center for Affective Sciences, University of Geneva, Campus Biotech, 9 Ch. Des Mines, Geneva 1202, Switzerland

vocalizations and facial expressions, showing the strong connection between the different expressive modalities (Hawk, Fischer, & Van Kleef, 2012).

There is some debate as to whether each modality carries specific information in the context of multimodal expressions. Researchers have used different approaches to investigate this issue, by focusing, for example, on how the synchronization between multiple signals has a facilitation effect in the interaction between mothers and children (Gogate, Bahrick, & Watson, 2000) and on the relative contribution of each modality when they provide conflicting (Massaro & Egan, 1996) or congruent emotional information (Bänziger, Grandjean, & Scherer, 2009; Martinez, Falvello, Aviezer, & Todorov, 2015; Wallbott & Scherer, 1986). While it seems clear and reasonable that multimodal expressions would be better recognized than unimodal expressions, the specific role of each modality remains to be determined (Martinez et al., 2015; Paulmann & Pell, 2011).

An interesting observation formulated in a recent study states that, while facial expressions seem to clearly differentiate negative emotions, a number of postural movements appear to differentiate positive emotion (Campos, Shiota, Keltner, Gonzaga, & Goetz, 2013). The results of one recent, detailed empirical analysis of bodily behaviour in 12 non-verbal expressions of emotion provide indirect support for this observation (Dael, Mortillaro, & Scherer, 2012). Close inspection of the results of the cluster analysis in the study reveals that the six positive emotions were more clearly differentiated by their bodily expression than the negative emotions were. For example, expressions of amusement formed a fully separated cluster while, more generally, the majority of emotion expressions for each positive emotion tended to fall within one cluster that did not include other positive emotions. Other studies have provided evidence of distinctive gestural signatures of pride (Tracy & Robins, 2004), love (Keltner, 1995) and amusement (Gonzaga, Turner, Keltner, Campos, & Altemus, 2006).

The emotion of interest (Izard, 1994; Silvia, 2006) provides an excellent opportunity to investigate the relative contribution of different modalities in emotion recognition and to explore the possibility that positive emotions may be differentiated more easily in terms of bodily movements than by facial expressions. Interest was in fact originally included in the facial expressions that Ekman tested for universality before it was eventually rejected from the final list of six emotions (Ekman, 1993), presumably on the grounds that it could not be reliably recognized facially. While the discovery of a prototypical facial expression of interest has so far proved elusive, emotion recognition studies indicate that it might be possible to recognize interest from dynamic multimodal expressions (Bänziger, Mortillaro, & Scherer, 2012). Furthermore, there are theoretical arguments that interest may have a very important evolutionary function that would justify the hypothesis that its expression should be consistently expressed and recognizable (Clément & Dukes, 2013). In light of this, we have carried out an exploratory study to investigate whether interest is more easily recognized when the whole body can be taken into account, rather than the facial expression alone. Additionally, we have examined whether dynamic presentations of expressions of interest are more readily recognizable than static images.

The facial expression of interest[1]

To our knowledge, the first major study of emotion recognition including interest was carried out over 50 years ago (Tomkins & Mc Carter, 1964). In their study, the authors adopted a forced-choice paradigm asking observers to recognize affects based on still head photographs of children and adults. Participants in the study were told to consider each one of the stated emotions as an example of a family of emotions. Each of the six traditional basic emotions is found in its "own" family, while Tomkins and McCarter also included interest ("excited", "attentive", "alert"), neutral ("casual", "emotionless", "without feeling") and ashamed ("guilty", "shy", "discouraged", "feeling inferior"). Results showed that, in descriptive terms, the expression of interest was the least often recognized of all the emotion categories. Izard carried out a series of studies in 1971, that he more fully described in 1994 (Izard, 1971, 1994), the most relevant of which was carried out with a procedure similar to that of Tomkins and McCarter's earlier research. Participants were split into two groups and the participants in each group were asked to place 32 photographs of facial expressions into one of eight emotion categories: interest, joy, surprise, sadness, anger, disgust/contempt, fear, shyness/shame. There were four photographs per category.

Each category was defined by a low and high intensity descriptor (e.g., anger-rage, interest-excitement), followed by four synonyms of the basic category. Thus, participants were being asked to match the photographs to the different discrete emotions. Only sadness and shyness/shame were recognized (in descriptive terms) less frequently than interest. The results support the findings of the earlier study, positioning interest as comparatively more difficult to recognize than other emotions.

Reeve (1993) carried out a study with adults in which five facial behaviours (eyes closed, number of eye glances, duration of eye glances, eyelid widening, exposed eyeball surface) and two head movements (head turns, head stillness) were associated to self-reports of interest while participants watched short films. In a later study designed to investigate intrinsic motivation, Reeve and Nix (1997) investigated the extent to which five acts of exploration (latency to initiate, comprehensiveness of investigation, richness of investigation, hand speed, approach-avoidance) and four facial displays of interest chosen from their earlier study (eye contact, eyeball exposure, eyes closed, lips parted) corresponded to self-reports of interest (and other measures) in 60 undergraduates while playing with a spatial-relations puzzle. They concluded that three of their exploration measures were indicative of intrinsic motivation, while only the time that the participants spent looking at the puzzles was indicative of interest. In terms of facial displays, they showed that there are (at least) two identifiable but different facial displays of interest that correlated with the self reports (namely eye contact, positively, and eyes closed, negatively). The authors concluded that a coherent cluster of facial interest displays does not exist, given that the two facial displays did not correlate with each other (Reeve & Nix, 1997).

However, Izard and colleagues described an expression of interest in several studies on babies (Izard, 1979, as cited in Langsdorf, Izard, Rayias, & Hembree, 1983; Izard, 1980), settling for the description of the expression of interest in their infant affective MAX coding system as "Brows can either be raised, in normal shape, or drawn together and/or slightly lowered; eyes may be either widened, roundish in appearance, or squinted; cheeks may be raised; mouth can be opened and relaxed, tongue may be extended beyond the gum line, lips may be pursed" (Matias, Cohn, & Ross, 1989, p. 489). Unlike Ekman and Reeve, Izard's work with children suggests that a facial expression of interest might exist, although, as the list indicates, the exact components of the expression are sometimes contradictory.

Rather than investigating whether certain static characteristics may or may not correlate with one another, another approach might be to assess dynamic representations of interest. As noted in one recent study: "it is not the simple presence of an AU that conveys certain information but rather the unfolding of an AU pattern with all its temporal characteristics" (Mortillaro, Mehu, & Scherer, 2011, p. 270). Indeed, the possibility of a dynamic expression of interest would mean that it could include several movements: it could well be that the eyebrows start in *normal shape* and are then subsequently *drawn together*, for example. Mortillaro et al. (2011) showed that it was indeed possible to distinguish interest from three other positive emotions (pride, pleasure, joy), in terms of frequency and duration of certain action units (AUs). The authors concluded that while "[s]everal features distinguish different pairs of emotions … no feature clearly sets one emotion apart from the others nor distinguishes pride from joy or interest from pleasure" (p. 268). Thus, they provide distinguishing features among the four positive emotions but no prototypical characteristics for each expression. Nonetheless, these results support the idea of interest as an unfolding expression – not one that can be distinguished statically. A more recent study (Campos et al., 2013) invited participants to pose expressions of different positive emotions after describing a situation when they had felt that emotion. While "brow lowering" (action unit (AU)4 from FACS; Ekman & Friesen, 1978) was analysed as being moderately to strongly associated with the expression of interest, none of the other AUs mentioned in their study were rated as being present in the Mortillaro et al. (2011) study. However, as far as we know, the facility with which the posed expressions of Campos et al. (2013) are recognized has not yet been studied.

The bodily expression of interest

Another important source of information for the recognition of interest may reside in its bodily expression. According to different theorists, interest itself includes an orienting response towards a novel stimulus, and this may be more apparent in the body than in the face. Camras et al. (2002) describe how "body stilling and facial sobering" characterize a reaction to a novel stimulus: the whole body and facial expression remains motionless for a moment after the stimulus becomes known. Scherer, Zentner, and Stern (2004) described this phenomenon as "freezing". This facial and body freezing has been described as being more likely to be found in babies and young infants, who may not have the cognitive capacity to disambiguate events and objects in their environment – as if they need to "buy more time" to appraise (Clément & Dukes, 2013). If this "freezing" is part of the expression of interest as a reaction to novelty (Silvia, 2005) or as a moment of attending and identification (Frijda, 1986), it could presumably be best captured in dynamic presentations and would be more likely captured when the whole body is presented rather than just the face.

Furthermore, given that we assume that being interested in something is synonymous with investing further in it, we could reasonably expect that this freezing is followed by a movement *towards* the object, in the same sense that being surprised, scared or disgusted by something may lead to taking a step back. Indeed, Kudoh and Matsumoto (1985) found that an expression of interest was positively related to "leaning forward" and "straightening one's back" while negatively to "turning one's back", "slowly turning one's head" and "turning one's head". Elsewhere, Frijda (1986) argues that the function of interest is "orientation", that its end state is "identification" and that it's action tendency is "attending (opening)". In comparison, in his study of the appraisal structure of interest, Silvia (2005) identifies interest as part reaction to a novel or ambiguous stimulus and part reaction to the potential to cope with the stimulus: if it is too complex or too simple, it is not interesting. Even if they differ slightly, these characterizations suggest that there may be a moment of approach towards, or at least of exploration of, something new or original after a moment of "reflection". Moreover, Reeve and Nix (1997) included *approach-avoidance* as a measure of exploration in intrinsic motivation, defining it as "the extent to which the participant drew near" or as "approach" (Reeve & Nix, 1997, p. 243). Interestingly, this measure correlated significantly with self-reported interest.

One of the few studies that include bodily expressions of interest uses the GEMEP corpus, a database of 18 emotions expressed in different modalities (Bänziger et al., 2012). One of the phenomena investigated by the authors relates to the patterns of bodily behaviour that differentiates emotions (Dael et al., 2012). The authors found that interest was expressed by a few sets of bodily movements that were uncommon in other positive emotions, and with one pattern exclusive to interest and irritation. The authors describe this cluster of movements: arms resting at side with trunk leaning frontal, combined with asymmetrical one arm action. They functionally interpreted this cluster as behaviours "represented in the action mode attending" (Dael et al., 2012, p. 1095), drawing on definitions made by Frijda as "wanting to observe well, to understand, to pay attention" (Frijda, 1986; Frijda, Kuipers, & Ter Schure, 1989). Results are very similar to those of a coding study by Wallbott (1998) in which two trained coders concluded that during actors' expressions of interest, "lateralized hands/arm movements" and "arms stretched out frontal" indicated hot anger, cold anger or interest in comparison to the other emotions.

Thus, interest appears to involve a moment of "freezing" followed by "approach towards", rather than avoidance of, the new stimulus (Roth & Cohen, 1986). This "approach towards' is likely either to take the form of leaning or moving towards, depending on the proximity of the object to the person, and its size. It may be possible then that there is a postural signature of interest that is not apparent in facial expressions.

In light of the absence of consensus about the existence of a prototypical expression, we decided to take an exploratory approach to investigate the expressive features of interest by considering the potential role of movement and the relative contribution of the body and the face to the recognition of interest compared to other emotions traditionally considered as having a prototypical facial expression (anger, disgust, fear, joy, sadness, surprise). In order to ensure the validity of our stimuli, comparison is made directly with how the actors portrayed interest facially in terms of previous studies (Mortillaro et al., 2011; Reeve, 1993) and the existing literature on corporally displayed interest is also used as a comparison (Dael et al., 2012; Wallbott, 1998).

We have taken two steps to make our analysis as clear as possible. The first is to call the static facial condition "the reference condition"; comparing results to it as a sort of baseline condition. Secondly, we will focus on a comparison between interest and the other emotions combined together (anger, disgust, fear, joy, sadness, surprise). Specifically, we predicted that (1) interest would be better recognized when presented dynamically rather than statically and (2) that it would be better recognized in the whole-body conditions rather than in the face conditions. Furthermore, we hypothesized that the recognition rate of interest will improve more than that of the composite recognition rate of the other emotions when presented dynamically (dynamic face vs. static face) and in whole-body form (static body vs. static face) rather than in the reference condition.

Moreover, and based primarily on Tomkins and Mc Carter's (1964) results and the afore-mentioned presumptions of Ekman's work, we predicted that in the reference condition (static face), interest would be the least recognized of the expressions. As a comparison, we predicted that there would be no difference between the recognition rates of interest and the composite recognition rates for the other emotions in the dynamic and whole-body conditions: not only would the comparative position of interest improve in regard to the other emotions, but interest would also become as recognizable as the other emotions combined together.

Method

Participants

A total of 125 adult volunteers (72 female) were recruited through student groups and by the use of social media. The participants were all French speaking. The average age of participants was 25 years, 10 months ($SD = 5$ years, 8 months), range = 19–44 years old. All had normal or corrected-to-normal vision. Sample size was established before beginning data collection and was based on our experience in previous similar perceptual studies.[2]

Stimuli and apparatus

The stimuli were presented to the participants using online survey software (Qualtrics). Given the lack of an available prototype for the expression of interest, it was impossible to instruct actors how to pose the expressions to create the interest stimuli. As a direct consequence, and with a view to using the same method for all the stimuli, we employed a technique for producing the stimuli that has been used elsewhere (for example, in creating the GEMEP corpus; Bänziger & Scherer, 2010; Bänziger et al., 2012). A film director, two actresses and studio space were hired. The actresses wore their hair tied back and a simple black round-neck t-shirt and denim jeans. The filming took place in a photograph studio, against an off-white background. The film director was present throughout the process.

As we were primarily concerned with how easily the expression of interest is recognized as a signal of investment in an object or event by a third party, the actresses posed with their eyes averted, as though they were looking at something off-camera. In the dynamic conditions, we asked the actresses to look around for a moment before fixing the target and reacting to it. Similar techniques for eye-gaze have been used in other studies (e.g., Conty & Grèzes, 2012).

For the face stimuli, the actresses were seated and were filmed from their shoulders up. They were instructed to keep their shoulders still and to look to the left or the right, depending on the condition and to act as though they were looking around aimlessly in the general direction of a fixed point on the wall (found at approximately 45 degrees), before fixing the target itself. The actresses were then asked to express the required emotion as a reaction to the target. While they were free to vocalize during their performance, neither of them did.

For the whole-body stimuli, the actresses were filmed from their knees to above their heads, walking to the centre of the shot. Again, they were asked to look aimlessly in the direction of the point on the wall. They were instructed to stop in the centre of the shot as they fixed on the point. Once standing in the centre of the screen, they were instructed to use their whole body to express the particular emotion elicited by the point.

Films of a duration of 6 s were taken for each emotion. For the static conditions, images were chosen from the film when the expressions appeared at their apex. Thus, the static face pictures were the maximum representation of the dynamic face stimuli and the static body pictures were the maximum representation of the dynamic body stimuli. The photos were all sized 1920 × 1080 with a resolution of 96 ppi. Examples of the static interest stimuli are provided in Figure 1.

The full list of emotions portrayed were anger, disgust, fear, interest, joy, sadness and surprise, while a neutral expression was included for comparison. The order in which the emotions were expressed was randomized to avoid anticipation in the actresses' portrayal of each emotion. To ensure that the emotions were expressed as ecologically as possible, once the experimental design had been explained to them, the actresses were encouraged to interpret the emotions as naturally as they could. This means that actresses were not instructed or trained to express emotions following Ekman's prototypes, but were free to express each emotion in their most spontaneous way.[3] The film director worked with each actress throughout the filming, ensuring that they were both comfortable with their portrayals.

Research design

Participants were proposed a within-subjects design: each participant was asked to recognize all of the stimuli. The first independent variable was EMOTION (interest vs. "other emotions"), which compared interest against all other emotions combined. The second independent variable was MOTION (static vs. dynamic), by which it was tested whether the stimuli were more easily recognized from being presented dynamically (as films) or statically (as photographs), irrespective of whether the face or whole-body conditions were presented. The third independent variable was that of MODALITY (face vs. whole body), by which it was established whether there was an advantage in terms of the recognition rates of the expression of interest being presented as "whole body and face" rather than "shoulders and face", irrespective of whether the stimuli were static or dynamic.

Procedure

The stimuli were split into two blocks: the two static conditions and the two dynamic conditions. The order of presentation of each block was randomized across the participants, and the order of the stimuli was randomized within each block. It was possible for the participants to take a break between the two blocks if they felt they needed one. If carried out without a break, the experiment took 25–35 minutes to complete.

Each stimulus was shown for 6 s, whether for the entirety of the film or the still photograph. The participants reported their response by selecting the emotion that they felt best described the actress's expression. There was no time restriction.

Apart from the affective responses, participants were instructed to select "other emotion" if they felt

Figure 1. Examples of stimuli. Top: details of the static whole-body interest stimuli. Bottom: examples of the static head stimuli.

that the actress was displaying an emotion that did not feature on the list. The full randomized list then, was composed of anger (*colère* in French), disgust (*dégoût*), fear (*peur*), interest (*intérêt*), joy (*joie*), neutral (*neutre*), other emotion (*autre emotion*), sadness (*tristesse*) and surprise (*surprise*).

Stimuli validation

The four static expressions of facial interest and the four dynamic expressions of facial interest were coded using the Facial Action Coding System (FACS) by three certified coders (Ekman & Friesen, 1978). When there was a discrepancy, then coders tried to resolve it; when impossible, we accepted the code of the senior coder who has more than 30 years of experience. The percentage of AUs observed in the static and dynamic facial expressions are reported in Table 1. Table 1 also includes the percentages of the AUs that were present in at least 40% of the expressions of interest in the most recent reference study in order to allow a direct comparison with our stimuli (Mortillaro et al., 2011).

The results in the current study are in line with the results from the 2011 study. The only action unit observed more than 50% of the time for expressions of interest in Mortillaro et al. (2011) that doesn't feature in the coding of the current expressions of this study is AU17 (chin raiser, 70%). This difference can be easily explained by the fact that, in the previous study, the actors were speaking and this may have increased the frequency of certain AUs, including AU17. With the marginal exception of Nostril Dilation (AU38), all the AUs coded as present in the current stimuli were listed as having been present in at least 40% of the interest expressions in Mortillaro et al. (2011). Furthermore, as can be seen from Table 1, there is more homogeneity in the dynamic stimuli: both AU4 and AU7 are coded as present in all the expressions. However, that these AUs did not co-occur (they were not displayed at the same point in time) seems to indicate the possibility of a dynamic sequence of movements for the expression of interest. This possibility is supported by the fact that AU1 and AU2 are present only in the dynamic stimuli suggesting that the unfolding of the expression of interest cannot be captured in a single frame.

Results

Table 2 shows the recognition rates and the false attributions for the static and dynamic facial and whole-body stimuli of interest. It is important to remember that the static stimuli were stills taken from the dynamic films.

The analysis of the incorrect answers shows that the participants were most likely to mislabel the expressions of facial interest for "other". Interestingly, facial expressions of interest were virtually never confused for expressions of joy. It also appears that the participants were increasingly likely to choose surprise as the amount of perceptual information increased,

Table 1. Presence of facial action units in the facial expressions of interest.

	AU1 Inner brow raiser	AU2 Outer brow raiser	AU4 Brow lowerer	AU7 Lid tightener	AU10 Upper lip raiser	AU12 Lip corner puller	AU17 Chin raiser	AU25 Lips parted	AU26 Jaw drop	AU38 Nostril dilator
Mortillaro et al. (2011)	.40	.50	.50	.90	.60	.80	.70	1	.50	.00
Static Faces (Pictures)	.00	.00	.50	.75	.50	.50	.00	.50	.50	.00
Dynamic Faces (Videos)	.50	.50	1	1	.50	.50	.00	.50	.50	.25

Note. The numbers represent the proportion of occurrence of AUs in the respective category of stimuli. The first row includes the data about the presence of these AUs in the portrayals of interest used in the study by Mortillaro et al. (2011). The second and third rows refer to the stimuli used in the present study.

Table 2. Recognition rates and false attributions for stimuli of interest.

	Interest	Anger	Disgust	Fear	Joy	Neutral	Sadness	Surprise	Other
Static Face	.29	.04	.12	.05	.03	.11	.12	.02	.22
Dynamic Face	.55	.01	.07	.02	.01	.10	.02	.05	.18
Static Body	.62	.06	.01	.03	0	.08	.04	.12	.08
Dynamic Body	.75	.01	.01	.01	0	.01	0	.20	.03

Note. Table 2 shows the average proportion of attributions for the four categories of stimuli of interest. Table 2 includes the falsely attributed expressions – for example, the Anger column shows how many times the participants erroneously chose anger when presented with a stimulus of an expression of interest.

Table 3. Recognition rates across the four conditions for all the expressions.

	Interest	Anger	Disgust	Joy	Neutral	Fear	Surprise	Sadness	Composite Recognition Rate (excluding interest)
Static Face	29 (144)	50 (251)	88 (441)	93 (465)	60 (302)	33 (167)	77 (386)	73 (365)	68
Static Body	62 (310)	56 (281)	63 (313)	71 (356)	81 (406)	55 (276)	88 (439)	41 (203)	65
Dynamic Face	55 (276)	54 (286)	78 (389)	85 (424)	76 (380)	62 (308)	80 (400)	81 (403)	74
Dynamic Body	75 (377)	51 (254)	61 (307)	62 (309)	82 (408)	68 (340)	86 (429)	48 (238)	65

Note. Scores are given in percentage (and raw) terms. The column "Composite Recognition rate" refers to the composite average score (in percentage terms) for all the expressions (including neutral) except interest in each condition (calculated as the average of the recognition rates of the other emotions). This "Composite Recognition Rate" score was the score against which the interest scores were measured in raw terms.

from static face to dynamic face and from static body to dynamic body.[4]

The recognition rates for interest and the other emotions combined (composite recognition rate) were calculated and compared across the conditions (Table 3). Repeated Measures ANOVA revealed that there was a main effect of EMOTION, $F(1, 124) = 60.540$, $p < .001$, $\eta_p^2 = 0.328$, that there was a main effect of how the stimuli were presented, as either faces or whole bodies, MODALITY, $F(1, 124) = 51.796$, $p < .001$, $\eta_p^2 = 0.295$ and as either still images or videos, MOTION, $F(1, 124) = 61.165$, $p < .001$, $\eta_p^2 = 0.330$. The results are shown in Figure 2.

Further analysis revealed an interaction of EMOTION and MODALITY, $F(1, 124) = 195.915$, $p < .001$, $\eta_p^2 = 0.612$, an interaction of EMOTION and MOTION, $F(1,124) = 53.694$, $p < .001$, $\eta_p^2 = 0.302$ and an interaction between MODALITY and MOTION, $F(1, 124) = 14.559$, $p < .001$, $\eta_p^2 = 0.105$. There was also a three-way interaction between EMOTION, MODALITY and MOTION, $F(1, 124) = 4.694$, $p < .05$, $\eta_p^2 = 0.036$.

To follow up on this three-way interaction, we analysed the effects of MODALITY and MOTION on the interest stimuli across the four conditions. Repeated Measures ANOVA revealed that there was a main effect of MODALITY, $F(1, 124) = 127.49$, $p < .001$, $\eta_p^2 = 0.507$, and of MOTION, $F(1, 124) = 64.59$, $p < .001$, $\eta_p^2 = 0.342$. There was also an interaction between MODALITY and MOTION, $F(1,124) = 9.811$, $p < .001$, $\eta_p^2 = 0.073$. This means that the recognition rates for interest improved when participants were presented with whole body stimuli rather than face stimuli, and with dynamic stimuli rather than static stimuli, and this last effect was larger for faces than for whole body expressions.

To compare interest with the other emotions, we calculated the difference in recognition rates for both interest and the other emotions combined between the static face condition and the dynamic face condition, and between the static face condition and the static whole-body condition. These differences were then compared with a view to testing whether the gain was greater for interest than for

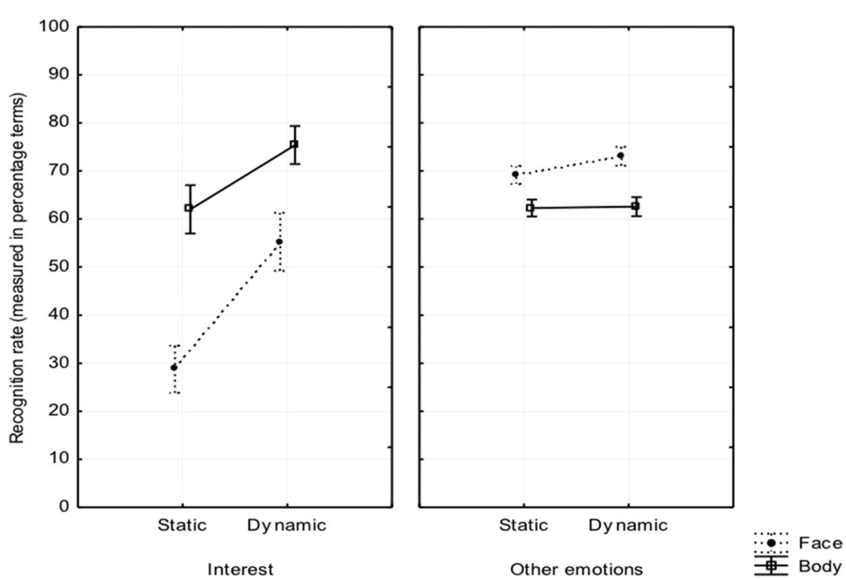

Figure 2. Recognition rate (measured in percentage) and confidence levels (95%). The average for interest is presented on the left and the composite average of the other emotions is presented on the right.

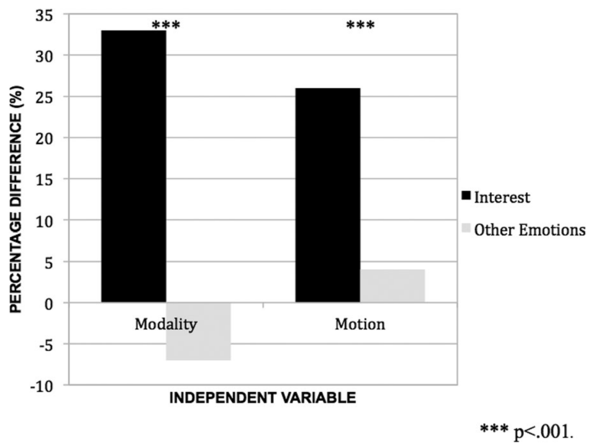

Figure 3. Difference between interest and the composite average of the other emotions in MODALITY (static body minus static face) and MOTION (dynamic face minus static face) in terms of percentage of recognition rate.

the combination of the other emotions. We analysed the relative gains for MODALITY and MOTION. The results are displayed in Figure 3 for both interest and the combination of the other emotions. It can be seen that the relative gain in recognition rate for interest in being presented either as a whole body or dynamically rather than in being presented in the reference condition is greater than the equivalent gain for the other emotions combined together. Planned contrasts with the reference condition revealed that the improvement in recognition rates for interest was greater than for the composite recognition rates of the other emotions when being presented in the whole body format (MODALITY), $t(124) = 11.59$, $p < .001$, and when being presented dynamically (MOTION), $t(124) = 6.39$, $p < .001$.

We compared the recognition rates for interest and the average recognition rates for the other emotions across all conditions: the other emotions were more easily recognized than interest in both the static face condition, $t(124) = 16.57$, $p < .001$, and the dynamic face condition, $t(124) = 5.96$, $p < .001$. There was no difference between the recognition rates of interest and the other emotions combined for the static body condition, and interest was more frequently recognized than the other emotions combined in the dynamic whole-body condition, $t(124) = 6.26$, $p < .001$.

Discussion

In real life, emotional expressions only rarely need to be recognized "instantaneously" and without any contextual cues – environmental, temporal, corporal, audible, etc. In light of this fact, combined with the new possibilities provided by technological progress, emotion recognition studies are progressively broadening their focus, considering videos in lieu of still pictures and multimodal expressions instead of facial expressions in isolation. One consequence of this broadening of focus has been an increase in the number of emotions considered in research on emotion recognition and in particular of positive emotions. Some of these studies led to the hypothesis that positive emotions may be reliably recognized and differentiated from one another when their expressions are presented dynamically and in whole-body form (Campos et al., 2013; Dael et al., 2012). Our results support this recent hypothesis by indicating that interest may be more easily recognized in perceptual conditions that are more ecological (with a view of the whole body and movements) when it is at least as likely to be recognized as the six more traditionally tested emotions combined together. It should be noted that by comparing interest against a composite average of emotions allows us to discuss the findings concerning interest but inevitably leads to masking the individual patterns of frequency with which each of those other emotions is recognized.

Given the results here and in earlier studies, it could be argued that no static facial prototype of the expression of interest exists, and that is why none has been found to date. Certainly, the dynamic presentation of the facial stimuli resulted in improved recognition rates, and this may indicate that it is the unfolding and the sequence of facial movements that makes recognising interest easier: maybe an expression of interest has two components, a reaction to novelty, reflected perhaps by AUs 1 and 2, and then a further commitment to investigation reflected perhaps by AUs 4 and 7. This is reflected in a changing expression.

As for the static presentation, our face stimuli were recognized around 30% of the time. Besides recognition rates, there are two other criteria to assess whether a stimulus is valid or not; the first criterion is to compare the expressions themselves with previous expressions in terms of action units. We found that our expressions are to a great extent in line with previous studies for which FACS codes of interest expressions are available. The second criterion, although more indirect, is to compare the recognition rates with earlier examples of stimuli purporting to represent the same expression.

But as stated earlier, Ekman abandoned his search for a universally recognizable expression of interest in static facial form (e.g., Ekman & Cordaro, 2011), Reeve and Nix doubted such an expression's existence (1997) and Mortillaro et al. said that there was not one feature that could distinguish interest from a set of positive emotions (2011). This would suggest, at the very least, that it is very difficult to recognize the static facial expression of interest, but, more controversially, that there may be no such thing as a prototypical static facial expression of interest.

However, Tomkins and Mc Carter (1964) found that interest was recognized about 58% of the time in their static face condition, while Izard (1971, 1994) recorded an average of 78% and 67% for his two studies over his four photographs. These figures are considerably higher than the 29% obtained here. Although the improved scores in those studies could probably be explained by the simple fact that they used photographs that were more readily recognizable as interest than the stimuli used here and in the previous studies mentioned, they could also be explained by the fact that their participants were presented with a list of a family of emotions, of which the named emotion was the principal one, rather than just the simple emotion-word as was used here; Tomkins and McCarter's method, also used by Izard, gave more context than is normally associated with emotion recognition paradigms, as each emotion category was more clearly elaborated. In the Tomkins & Mc Carter, 1964 study, for example, the interest group also had the labels "excited", "attentive" and "alert" attached to it. Any one photograph that looked like an expression of interest, excitement, attentiveness or alertness was chosen as belonging to that same category. It could be that the participants recognized a given expression as "alert" and, had "alert" not been explicitly stated as belonging to the "interest family", the participants may have been tempted to put the alert expression with "fear" or "surprise". The method of including families of expressions thus makes the target expression much "larger to hit" by the provision of some semantic context. Also, while Izard reported all of the results (albeit 23 years after the test was carried out), Tomkins and McCarter only reported the results of the three photographs for which there was the greatest consensus: had they presented the average results of the seven photographs with which their participants were presented, the reported score would necessarily have been lower.

In spite of these differences in experimental design, the comparative position of interest remained the same: as far as we are aware, when interest has been part of a static facial emotion recognition task, it has almost always fared less well than all of the more traditionally tested emotions featured here. With that in mind, we could argue that the static facial stimuli used for this study represent *fragments* of expressions of interest: only the provision of contextual information can permit the assembly of these fragments to be recognized as the full expression, that is, when they are presented dynamically or in whole body format. Reeve (1993, p. 373) came to the same conclusion, stating that "to portray the full face of interest, it seems necessary to supplement the [elements] which can be portrayed in a still photograph with an over time display … which can best be portrayed in a multiple-second videotape".

Whether a static facial prototype exists or not, our results show that presenting the stimuli in a dynamic format considerably increases the accuracy of recognition. This is in line with other recent studies that put forward the importance of dynamics for emotion perception (Bänziger et al., 2012; Krumhuber, Kappas, & Manstead, 2013; Nelson & Russell, 2014). In fact, our results indicate that the relative gain for interest when presented dynamically rather than statically was significantly higher than the one for the composite average of the other emotions.[5] This again suggests that there is probably something about the behavioural expression of interest that cannot be easily grasped with one picture. If we turn to appraisal theory and models, interest has been described as characterized by appraisals of novelty and coping potential (Silvia, 2006). Freezing is a response to novelty that aids sensory acuity and, if we are to incorporate Silvia's model of interest, could also buy time to judge how easy it is to "cope" with the stimulus. If the object is then appraised of being worthy of further attention, this freezing behaviour is followed by an approach response as interest develops, enabling exploration and investigation. For example, imagine a picture of someone with their eyebrows lowered (AU4) and the eyelids tightened (AU7). It might not be clear if the person is expressing interest or if they are squinting (because of bright sunshine, for example). However, if there were a short film of that same person freezing to look at something before "squinting", it would be much clearer that this person is reacting with interest to a stimulus rather than as a physiological reaction to light. The snapshot might look exactly the same, but the dynamic component of the film provides vital contextual information. In summary, someone may freeze to look at something because it captures his or her attention. By leaning in or approaching, the person is making a further commitment: the object is not just attention grabbing – it is interesting.

A recent study has shown that body cues rather than facial expression can be used to discriminate between intense positive and negative emotions (Aviezer et al., 2012). In 2004, Tracy and Robins briefly describe a study (in a footnote) where they set out to test whether pride could be recognized from the face alone (Tracy & Robins, 2004, p. 196). None of their facially expressed pride stimuli was recognized at greater than chance level. Their conclusion was that both the face and the body are necessary to recognize an expression of pride (although see Nelson and Russell, 2014, for evidence to the contrary). Similarly, in the present study, it was only when presented in whole body form that interest was (at least) as well recognized as the average of the other emotions. Our whole body stimuli have several similarities with the descriptions of body movements of interest given in previous studies (Dael et al., 2012; Wallbott, 1998): actors are either leaning their face or head and chest forward, while only one of them doesn't have their hands by their sides but on their legs before them. They also appear to have taken a small step forward. This could suggest that the extra information which is available to the participants by presenting an expression dynamically, rather than statically, exists already in the static whole-body stimuli, as if the static whole body stimuli includes dynamic information – with one foot forward, leaning forward (Dael et al., 2012). Perhaps counter-intuitively then, a static photograph of the whole body captures the "freezing" just as well as a short film.

The Ekman photographs have been used extensively in affective science and beyond as representations of basic emotions. The original idea was that these six emotions – anger, disgust, fear, joy, sadness, surprise – were universal signals of how the individual felt and that could be "read" by others to predict the behaviour of the expresser. Joy has traditionally been argued to be the only positive member of the basic emotion group but this claim was essentially only based on studies on facial expression.[6] Given the high score of most of the static body stimuli of interest in the present study, it is also more likely that a prototypical expression of interest may be found in a composite of body and face rather than facial expression alone. Thus, it may well be that interest follows the same pattern as pride. Future research is necessary to determine whether interest and other positive emotions more generally have a distinctive gestural signature and the importance of that gesture being dynamically presented.

If we accept the theory that emotions are essentially social signals, it would seem that we have to communicate positive emotion more "loudly" than negative emotion for them to be perceived, that is, that the communication requires the whole body and not just the face. Perhaps, from the perceiver's side, it has been more important to be able to differentiate negative emotions from one another: in evolutionary terms, if all the positive emotions meant that there was no immediate danger, perhaps one facial signal of "all clear" has been sufficient. While it is probable that it is of evolutionary significance to be able to immediately distinguish between fine-grained facial reactions of anger, fear, disgust and even sadness (in the case of a sudden loss), it doesn't make as much sense in such terms to be able to instantly distinguish between joy, pride, pleasure and interest, as they are much less likely to need an immediate and vital response. For example, even though being able to recognize interest in a third-party may well be vital for successful living, in terms of being aware of accepted behaviour and valued items in any particular culture, it can hardly be described as being "immediate and vital" to detect an expression of interest (Clément & Dukes, 2013). It would follow, then, that no such static facial signal of interest (or, for example, pride or joy) has either developed or is recognizable. Nevertheless, the very fact that such positive emotions can be distinguished through facial expression and context, or a combination of facial and body expression, underlies their importance for general well-being.

The current study was intended to highlight the new and increasing number of studies dedicated to looking beyond the static face modality when considering emotion expression. The empirical component of this study is exploratory in its nature, with only two actresses and a limited number of stimuli. Furthermore, the results would have been more readily comparable across conditions had only one recording shoot been used; as it stands, it is possible that the facial expressions were more (or less) recognizable in one of the two shoots. Nevertheless, it appears to confirm that the static facial expression of interest is very difficult to recognize while hinting at the possibility that dynamics and bodily movements can make interest as easily recognizable as the other more traditionally used emotions. This study contributes to the growing body of literature that argues for a more global study of emotion expressions in which still photographs of faces are only one of the many options available to investigate the emotion expression and perception processes and, in some cases, it is not the most appropriate one.

Notes

1. Although we focus on facial and bodily expressions of interest, it is intriguing to note that in a vocal emotion recognition test of 14 emotions, Banse and Scherer (1996) found interest to be the third most recognized emotion, scoring an impressive 75% behind just happiness (78%) and boredom (76%).
2. A replication study was carried out using a between-subjects design and 27 participants for each of the four conditions. The results of that study are strikingly similar to the results reported here and therefore no further comment will be made about the replication study in this text. The results of the replication study are available on request from the corresponding author.
3. While the actresses were free to express the emotions as they wished, it is clear that the expressions are still not "authentic". As such, it may be that real-life expressions of interest would bear different results.
4. This was borne out by statistical analysis. We recoded the wrong answers by attributing a "1" to the trials where participants answered "surprise" and a "0" for any other answers. A logistic mixed effect model was performed on this variable in order to test how the probability to incorrectly answer "surprise" rather any other emotion was modulated by the factors of our design; the fixed effects we entered were MOTION and MODALITY in the model and the random effect we entered was the participants. Results showed that the participants were significantly more likely to answer surprise when they made a mistake in the dynamic face condition rather than in the face condition ($z = 3.487$, $p < .001$) and in the dynamic body condition rather than in the static body condition ($z = 1.471$, $p < .05$).
5. It could be that a dynamic presentation of the expressions did not affect the recognition of all other emotions the same way. For example, from a qualitative analysis of Table 3, it seems that the facial expressions of anger, fear, surprise and sadness benefit from being presented dynamically, while the facial expressions of disgust and joy do not. Meanwhile, in the body conditions, only fear and joy seem to change particularly when represented dynamically – joy is less well recognized while fear is recognized more frequently. Clearly it would be interesting to speculate about the reasons for the specific patterns uncovered here, but testing of these speculations and hypotheses is beyond the scope of the current manuscript. Furthermore, to carry out such analysis would entail the creation of many more stimuli and require testing that would be statistically inappropriate in our experimental design.
6. Surprise has sometimes been argued to be a second positive emotion, but see Noordewier and Breugelmans (2013) for a convincing explanation of why surprise is initially more likely to be negative than positive.

Acknowledgements

The authors would like to thank the participants and students who volunteered for this study and the replication study, the precious help of Nadège Foudon, the two actresses (Pauline Schneider and Sarah-Lise Salomon Maufroy), the director (Paul Walther) and the FACS coders who coded the facial expressions used in this study, Cyrielle Chapuis, Benoît Cano Peuker and Susanne Kaiser.

Disclosure Statement

No potential conflict of interest was reported by the authors.

Funding

The data collection was carried out at the University of Neuchâtel, Neuchâtel, Switzerland. This work was supported by the National Centre of Competence in Research (NCCR) "Affective Sciences: Emotion in individual Behaviour and Social Processes", financed by the Swiss National Science Foundation [grant number: SNSF, 51NF40-104897], and hosted by the University of Geneva.

ORCID

Catherine Audrin http://orcid.org/0000-0003-2905-6000

References

Aviezer, H., Hassin, R. R., Ryan, J., Grady, C., Susskind, J., Anderson, A., … Bentin, S. (2008). Angry, disgusted, or afraid? Studies on the malleability of emotion perception: Research article. *Psychological Science*, *19*(7), 724–732.

Aviezer, H., Trope, Y., & Todorov, A. (2012). Body cues, not facial expressions, discriminate between intense positive and negative emotions. *Science*, *338*(6111), 1225–1229.

Banse, R., & Scherer, K. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, *70*(3), 614–636.

Bänziger, T., Grandjean, D., & Scherer, K. R. (2009). Emotion recognition from expressions in face, voice, and body: The multimodal emotion recognition test (MERT). *Emotion (Washington, D.C.)*, *9*(5), 691–704. http://doi.org/10.1037/a0017088

Bänziger, T., Mortillaro, M., & Scherer, K. R. (2012). Introducing the geneva multimodal expression corpus for experimental research on emotion perception. *Emotion*, *12*(5), 1161–1179.

Bänziger, T., & Scherer, K. R. (2010). Introducing the Geneva multimodal emotion portrayal (GEMEP) corpus . In T. Bänziger, K. Scherer, & E. Roesch (Eds.), *Blueprint for affective computing: A sourcebook* (pp. 271–294). Oxford, England: Oxford University Press.

Campos, B., Shiota, M. N., Keltner, D., Gonzaga, G. C., & Goetz, J. L. (2013). What is shared, what is different? Core relational themes and expressive displays of eight positive emotions. *Cognition & Emotion*, *27*(1), 37–52.

Camras, L., Meng, Z., Ujiie, T., Dharamsi, S., Miyake, K., Oster, H., … Campos, J. (2002). Observing emotion in infants: Facial expression, body behaviour, and rater judgements of responses to an expectancy-violating event. *Emotion*, *2*, 179–193.

Clément, F., & Dukes, D. (2013). The role of interest in the transmission of social values. *Frontiers in Psychology*, *4*, 349.

Conty, L., & Grèzes, J. (2012). Look at me, I'll remember you: The perception of self-relevant social cues enhances memory and right hippocampal activity. *Human Brain Mapping*, *33*(10), 2428–2440.

Dael, N., Mortillaro, M., & Scherer, K. R. (2012). Emotion expression in body action and posture. *Emotion*, *12*(5), 1085–1101.

de Gelder, B., & Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cognition and Emotion*, *14*(3), 289–311.

Ekman, P. (1993). Facial expression and emotion. *American Psychologist*, *48*, 384–392.

Ekman, P., & Cordaro, D. (2011). What is meant by calling emotions basic. *Emotion Review*, *3*(4), 364–370.

Ekman, P., & Friesen, W. V. (1978). *Facial action coding system*. Palo Alto, CA: Consulting Psychologists Press.

Frijda, N. H. (1986). *The emotions*. Cambridge, Paris: Cambridge University Press and the Editions de la Maison des Sciences de l'Homme.

Frijda, N. H., Kuipers, P., & Ter Schure, E. (1989). Relations among emotion, appraisal, and emotional action readiness. *Journal of Personality and Social Psychology*, *57*(2), 212–228.

Gogate, L. J., Bahrick, L. E., & Watson, J. D. (2000). A study of multimodal motherese: The role of temporal synchrony between verbal labels and gestures. *Child Development*, *71*(4), 878–894.

Gonzaga, G. C., Turner, R. A., Keltner, D., Campos, B., & Altemus, M. (2006). Romantic love and sexual desire in close relationships.. *Emotion*, *6*, 163–179.

Hawk, S. T., Fischer, A. H., & Van Kleef, G. a. (2012). Face the noise: Embodied responses to nonverbal vocalizations of discrete emotions. *Journal of Personality and Social Psychology*, *102*(4), 796–814.

Izard, C. E. (1971). *The face of emotion*. New York: Appleton-Century-Crofts.

Izard, C. E. (1979). *The maximally discriminative facial movement coding system (Max)*. Newark, DE: Instructional Resources Center, University of Delaware.

Izard, C. E. (1980). The emergence of emotions and the development of consciousness in infancy. In J. M. Davidson & R. J. Davidson (Eds.), *The psychobiology of consciousness in infancy* (pp. 193–215). New York: Plenum Press.

Izard, C. E. (1994). Innate and universal facial expressions: Evidence from developmental and cross-cultural research. *Psychological Bulletin*, *115*(2), 288–299.

Keltner, D. (1995). Signs of appeasement: Evidence for the distinct displays of embarrassment, amusement and shame. *Journal of Personality and Social Psychology*, *68*, 441–454.

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

Kudoh, T., & Matsumoto, D. (1985). Cross-cultural examination of the semantic dimensions of body postures. *Journal of Personality and Social Psychology*, *48*(6), 1440–1446.

Langsdorf, P., Izard, C. E., Rayias, M., & Hembree, E. A. (1983). Interest expression, visual fixation, and heart rate changes in 2-and 8-month-old infants. *Developmental Psychology*, *19*(3), 375–386.

Martinez, L., Falvello, V. B., Aviezer, H., & Todorov, A. (2015). Contributions of facial expressions and body language to the rapid perception of dynamic emotions. *Cognition & Emotion*, *9931*(May), 1–14.

Massaro, D. W., & Egan, P. B. (1996). Perceiving affect from the voice and the face. *Psychonomic Bulletin & Review*, *3*(2), 215–221.

Matias, R., Cohn, J. F., & Ross, S. (1989). A comparison of two systems that code infant affective expression. *Developmental Psychology*, *25*(4), 483–489.

McHugh, J. E., Kearney, G., Rice, H., & Newell, F. N. (2012). The sound of the crowd: Auditory information modulates the perceived emotion of a crowd based on bodily expressions. *Emotion (Washington, D.C.)*, *12*(1), 120–131.

Mondloch, C. J. (2012). Sad or fearful? The influence of body posture on adults' and children's perception of facial displays of emotion. *Journal of Experimental Child Psychology*, *111*(2), 180–196.

Mondloch, C. J., Horner, M., & Mian, J. (2013). Wide eyes and drooping arms: Adult-like congruency effects emerge early in the development of sensitivity to emotional faces and body postures. *Journal of Experimental Child Psychology*, *114*(2), 203–216.

Mortillaro, M., Mehu, M., & Scherer, K. R. (2011). Subtly different positive emotions can be distinguished by their facial expressions. *Social Psychological and Personality Science*, *2*(3), 262–271.

Mortillaro, M., Mehu, M., & Scherer, K. R. (2013). The evolutionary origin of multimodal synchronization and emotional expression. In E. Altenmüller, S. Schmidt, & E. Zimmermann (Eds.), *Evolution of emotional communication: From sounds* *in nonhuman mammals to speech and music in man* (pp. 3–25). New York: Oxford University Press.

Nelson, N. L., & Russell, J. A. (2011). Preschoolers' use of dynamic facial, bodily, and vocal cues to emotion. *Journal of Experimental Child Psychology*, *110*(1), 52–61.

Nelson, N. L., & Russell, J. A. (2014). Dynamic facial expressions allow differentiation of displays intended to convey positive and hubristic pride. *Emotion*, *14*, 857–864.

Noordewier, M. K., & Breugelmans, S. M. (2013). On the valence of surprise. *Cognition & Emotion*, *27*(7), 1326–1334.

Paulmann, S., & Pell, M. D. (2011). Is there an advantage for recognizing multi-modal emotional stimuli? *Motivation and Emotion*, *35*(2), 192–201.

Reeve, J. (1993). The face of interest. *Motivation and Emotion*, *17*(4), 353–375.

Reeve, J., & Nix, G. (1997). Expressing intrinsic motivation through acts of exploration and facial displays of interest. *Motivation and Emotion*, *21*(3), 237–250.

Roth, S., & Cohen, L. J. (1986). Approach, avoidance, and coping with stress. *American Psychologist*, *41*(7), 813–819.

Scherer, K. R., & Ellgring, H. (2007). Multimodal expression of emotion: Affect programs or componential appraisal patterns? *Emotion*, *7*(1), 158–171.

Scherer, K. R., Zentner, M. R., & Stern, D. (2004). Beyond surprise: The puzzle of infants' expressive reactions to expectancy violation. *Emotion*, *4*(4), 389–402.

Silvia, P. J. (2005). What is interesting? Exploring the appraisal structure of interest. *Emotion*, *5*(1), 89–102.

Silvia, P. J. (2006). *Exploring the psychology of interest*. New York: Oxford University Press.

Tomkins, S. S., & Mc Carter, R. (1964). What and where are the primary affects? Some evidence for a theory. *Perceptual and Motor Skills*, *18*(1), 119–158.

Tracy, J. L., & Robins, R. W. (2004). Show your pride: Evidence for a discrete emotion expression. *Psychological Science*, *15*(3), 194–197.

Van den Stock, J., Righart, R., & de Gelder, B. (2007). Body expressions influence recognition of emotions in the face and voice. *Emotion*, *7*(3), 487–494.

Wallbott, H. G. (1998). Bodily expression of emotion. *European Journal of Social Psychology*, *28*(6), 879–896.

Wallbott, H. G., & Scherer, K. R. (1986). Cues and channels in emotion recognition. *Journal of Personality and Social Psychology*, *51*(4), 690–699.

Watson, R., Latinus, M., Noguchi, T., Garrod, O., Crabbe, F., & Belin, P. (2014). Crossmodal adaptation in right posterior superior temporal sulcus during face–voice emotional integration. *Journal of Neuroscience*, *34*(20), 6813–6821.

Wenzler, S., Levine, S., van Dick, R., Oertel-Knöchel, V., & Aviezer, H. (2016). Beyond pleasure and pain: Facial expression ambiguity in adults and children during intense situations. *Emotion*, *16*(6), 807–814.

Zieber, N., Kangas, A., Hock, A., & Bhatt, R. S. (2014). The development of intermodal emotion perception from bodies and voices. *Journal of Experimental Child Psychology*, *126*, 68–79.