Exp Brain Res
DOI 10.1007/s00221-009-2037-5

RESEARCH ARTICLE

Perceiving emotion in crowds: the role of dynamic body postures on the perception of emotion in crowded scenes

Joanna Edel McHugh · Rachel McDonnell · Carol O'Sullivan · Fiona N. Newell

Received: 7 May 2009 / Accepted: 24 September 2009

Abstract Although the perception of emotion in individuals is an important social skill, very little is known about how emotion is determined from a crowd of individuals. We investigated the perception of emotion in scenes of crowds populated by dynamic characters each expressing an emotion. Facial expressions were masked in these characters and emotion was conveyed using body motion and posture only. We systematically varied the proportion of characters in each scene depicting one of two emotions and participants were required to categorise the overall emotion of the crowd. In Experiment 1, we found that the perception of emotions in a crowd is efficient even with relatively brief exposures of the crowd stimuli. Furthermore, the emotion of a crowd was generally determined by the relative proportions of characters conveying it, although we also found that some emotions dominated perception. In Experiment 2, we found that an increase in crowd size was not associated with a relative decrease in the efficiency with which the emotion was categorised. Our findings suggest that body motion is an important social cue in perceiving the emotion of crowds and have implications for our understanding of how we perceive social information from groups.

J. E. McHugh · F. N. Newell (✉)
School of Psychology, Institute of Neuroscience,
Trinity College Dublin, Lloyd Building,
Dublin 2, Ireland
e-mail: fiona.newell@tcd.ie

J. E. McHugh
e-mail: mchughje@tcd.ie

R. McDonnell · C. O'Sullivan
Graphics, Vision and Visualisation Group,
Trinity College Dublin, Dublin 2, Ireland

Keywords Social perception · Crowd scenes · Emotion

Introduction

Crowds of people are something we see on an everyday basis; from the queue at the bus stop, to hoards of shoppers on the streets. There is a wealth of social information which can be perceived directly from crowds. For instance, we may identify a group of football supporters on their way from a game by the clothes they wear, or we may identify a group of people shopping from the bags they carry. However, emotion is central to many aspects of social perception, and the perception of the behaviour of others helps to infer their intentions and traits (Allison et al. 2000). Indeed, it has previously been proposed that emotion is such an important part of perception that the first appraisal of our environment prioritises emotional information (Schindler et al. 2008). As such, the emotion of a crowd should be perceived directly in order for us to act appropriately. For example, the football supporters may be disappointed if their team lost or the shoppers may be excited by a bargain sale. It may therefore be important to rapidly perceive an emotional change in a crowd such as when disappointment turns to anger or when excitement turn to delight so that we either avoid or follow that crowd.

To date, most research investigating the perception of emotion has focused on facial expressions (for a review see Adolphs 2002). However, many recent studies have investigated the role of emotional body language (i.e. dynamic body movement) on the perceived emotional state of the character (see, e.g. de Gelder 2006 for a review). Indeed, it is proposed that facial, prosodic and body expressions provide similar information for the perception

Published online: 15 October 2009

of emotion in that the six basic emotions (anger, disgust, fear, sadness, happiness and surprise (Ekman 1992; Scherer 1981; Wallbott 1998) are readily recognisable from all three sources of information.

Evidence from studies on the perception of emotion from facial expressions has suggested that facial expressions are perceived as categorical (Young et al. 1997), i.e. that they are perceived as belonging to discrete categories irrespective of the visual similarity between items (see, Harnad 1987). Many studies have suggested that more complex visual stimuli such as faces and objects are perceived as categorical (Etcoff and Magee 1992; Newell and Bülthoff 2002) suggesting that it is a general mechanism for discriminating between similar events or items. Although Young et al. (1997) suggest that facial expressions are perceived as categorical, it is not known whether emotions from dynamic body expressions are also perceived as categorical. Facial expression is an effective cue for perceiving the emotional state of another individual (Ekman 1973, 1982), however, body expression is likely to be a more relevant cue for determining emotion when facial expressions might be difficult to resolve in individuals from afar (e.g. Walk and Walters 1988) or in a crowd of individuals. Considering the similarity between the perception of facial expression and of emotional body language, for example, they share the same electrophysiological signature (Stekelenburg and de Gelder 2004) and both are processed in a configural manner (Reed et al. 2003), it could be that emotions from body expressions are also perceived as categorical. Conversely, it could be that the direct perception of the relative proportions of characters conveying an emotion within a crowd is beneficial for behaviour as such perception allows for the efficient evaluation of the extent to which the crowd as a whole shares the emotion. For example, if only a small proportion of a crowd are angry and the majority are sad, it may be important to detect the angry characters within a crowd so that they can be avoided. If the emotion of crowds was perceived as categorical then it may not be possible to perceive the few angry characters.

In the present study, we investigated whether the emotion of crowds is perceived as categorical by systematically varying the relative proportions of characters expressing one emotion over another in a crowd. If these relative proportions were perceived as categorical then we expected that there would be a dramatic shift in the perceived emotional category of the crowd at the category boundary only and that the emotion of crowd stimuli either side of this boundary would be consistently categorised as one emotion over another, irrespective of the proportion of characters conveying the emotion. On the other hand, if the emotion of a crowd is not perceived as categorical then we would expect a systematic change in categorisation response with changes in the proportion of characters portraying an emotion.

Some research on emotional cognition has found that emotional cues can affect perception even when not consciously recognised (Whalen et al. 1998; De Gelder and Hadjikhani 2006; Morris et al. 2001; De Gelder et al. 1999; De Gelder et al. 2005). Other evidence has supported the idea that focussed attention may not be required to perceive the emotional category of an image. For example, the emotional 'gist' of a scene can be accurately perceived from a brief exposure to the scene of about 200 ms (Hampton et al. 1989). Such rapid exposure is too short for attention or eye movements to be deployed over the entire scene, suggesting that the emotional content of a scene is rapidly determined. Moreover, Hansen and Hansen (1988) found that some facial expressions can be rapidly detected in a crowd scene suggesting that the perception of emotion from faces at least may be preattentive (Treisman and Gelade 1980). If it were the case that focussed attention was not necessary for emotional cues to affect perception, then we might expect that all characters in a scene would affect perception. It is clearly important, not only for social interactions but often for survival, that the emotion expressed by a crowd is directly perceived. What is not known, however, is whether the accurate perception of emotion of crowds is rapid or requires attentional allocation to the body expressions of each of the individual characters, or a portion of the characters in a crowd, in order for the emotion to be determined. We addressed this question in the following experiments by investigating the effect of rapid exposure of dynamic crowd stimuli and the effect of number of characters on the perception of emotion in crowds.

In sum, the aim of our study was to investigate how the emotion of a crowd containing dynamic characters expressing emotion through body posture and movement is perceived. Crowds containing different proportions of emotional characters, and different overall numbers of characters, were created. It was expected that the perception of emotion in crowd scenes would be highly accurate and would be rapidly determined irrespective of stimulus duration and crowd size. However, it was unclear whether the emotion of a crowd would be perceived as categorical or not and our studies were aimed at elucidating this process.

Experiment 1

In the following experiment, we used dynamic scenes of crowds of characters expressing an emotion. For some of the crowd scenes all the characters conveyed the same, consistent emotion whereas for other crowd scenes some of

the characters conveyed one emotion whereas the other characters conveyed a different emotion. In other words, our crowd scenes depicted one of the two emotional categories and we systematically varied the proportion of characters depicting each emotion across scenes. Emotion was expressed in a scene through the dynamic motion of the characters' bodies only and facial information was masked. Virtual characters were used within the crowd scenes, as these characters offer more control over the use of the same emotional body language across different crowd scenes. Furthermore, previous studies have suggested that emotion is perceived as effectively from virtual characters as images of real individuals (Fabri et al. 2004; McDonnell et al. 2009).

We predicted that the emotion of a crowd would be more efficiently determined from a crowd of characters conveying a consistent emotion than crowds conveying more than one emotion. If the emotion of a crowd is perceived as categorical then for crowds containing characters with different emotions, we predicted that the emotion conveyed by either the larger proportion of the crowd, or the more visually salient emotion, would dominate the emotion perceived. Otherwise, if the perceived emotion of a crowd is determined by the relative proportion of characters displaying one emotion over another, then we would expect a monotonic increase in the perceived emotion with increasing proportions of characters conveying that emotion.

Methods

Participants

We recruited an opportune sample of 225 participants (117 females) who took part in a science exhibition at the Science Gallery, Trinity College Dublin. Their ages ranged from 6 to 78 years, with a mean age of 33.4. All participants gave informed, written consent to partake in the study. All reported normal or corrected-to-normal vision. The study (and all subsequent studies reported here) was approved by the Trinity College School of Psychology Research Ethics Committee.

Stimuli

Our stimuli were created by motion capturing the body movements of six actors who were required to demonstrate four different emotional states. All actors were required to use movement to display the emotions but to remain within a limited location during the actions. We used a Vicon IQ 2.0 optical system, with 10 infrared cameras and 41 markers placed according to the Vicon human template to motion capture the body movements. Actors were instructed to display, as naturally as possible, each of the following emotions using body language only: anger, fear, happiness and sadness. On average, we recorded 6 seconds of each emotion conveyed by each actor.

We then applied each of these motion captured files onto the body shapes of six virtual characters (3 male and 3 female) to create virtual characters with real human movement patterns (see, e.g. Sturman 1994). Characters were rendered with the motion capture files using 3D Studio Max 8 software, at a rate of 25 frames/s. Although six actors were originally recorded, the files of one female actor had to be discarded as artefacts appeared in the rendered image due to hand–head interactions during the motion capture. Each character was rendered using 10,000 polygons. Soft shadows and motion blur were incorporated to increase realism. In order to present the most informative dynamic cue related to each emotion we reduced the timing of each original motion file from 6 s to the best 2 s of footage determined on the basis of the results of a pilot study. To provide variation across our stimuli, each of the five virtual characters was rendered in five different viewpoints: forward facing; facing 30° to the left and right; facing 60° to the left and right. The actors (5), emotional (4) and viewpoint changes (5) yielded a total stimulus set of 100 dynamic, virtual figures.

Crowd stimuli were then created from this set of virtual characters using Adobe Premiere 6.5. Each crowd stimulus scene contained 10 male and 10 female figures. These images were rendered as greyscale images with virtual characters presented against a grey background (RGB value of 120, 120, 120). Each virtual figure was pseudorandomly positioned in each scene with the constraints that no character was placed at the centre of the crowd (where a fixation cross was positioned prior to the crowd stimulus), the body of each character was fully visible and no character occluded information from another character during the video sequence. Characters were randomly positioned in each scene across trials. Figure 1 provides an example of the crowd stimuli used in the task.

We created four basic crowd stimuli where all characters consistently depicted one of the four emotional categories. These stimuli represented the "100%" crowd stimuli. We then created continua between all possible pairings of these 100% stimuli by adopting the design previously described by Young et al. (1997). There were six different pairings between the stimuli: anger–fear, anger–happy, anger–sad, fear–happy, fear–sad, happy–sad. In turn, there were five steps within each continuum: a crowd scene where 100% of the characters depicted emotion 'A' and 0% emotion 'B'; a crowd scene where 75% depicted emotion 'A' and 25% emotion 'B'; a scene were 50% depicted emotion 'A' and the other 50% emotion 'B'; a scene where 25% depicted emotion 'A' and 75% emotion 'B' and finally a

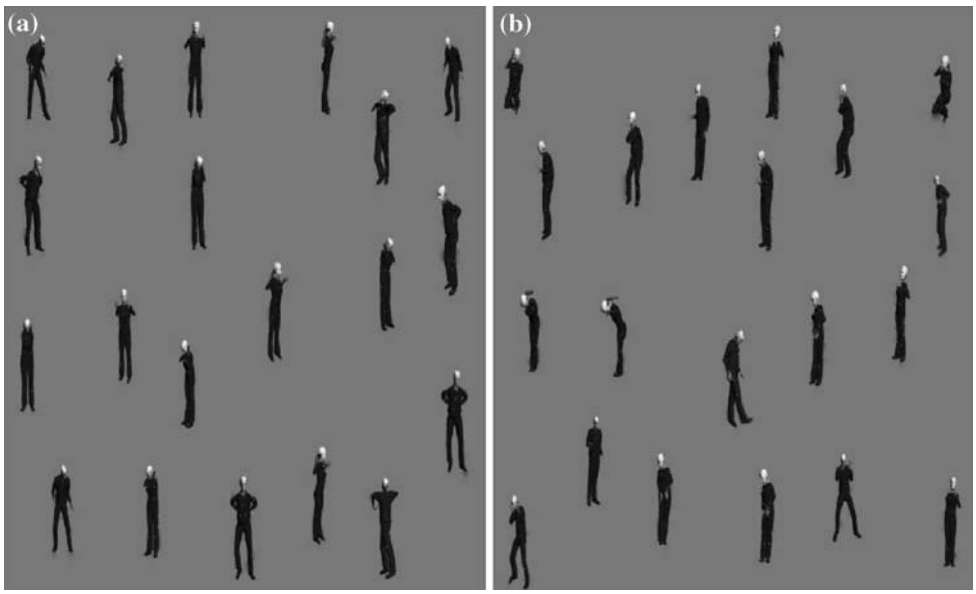

Fig. 1 Image examples of the crowd stimuli used in Experiments 1 and 2. Image (**a**) depicts a 100% consistent crowd of characters conveying 'anger' whereas image (**b**) depicts a crowd with 50% of the characters conveying 'sad' and 50% conveying 'fear'

scene where 0% depicted emotion 'A' and 100% emotion 'B'. For instance, along the anger–happy continuum, five crowds contained the following proportion of characters depicting each of the anger or happy emotions: 100% anger + 0% happy, 75% anger + 25% happy, 50% anger + 50% happy, 25% anger + 75% happy, 0% anger + 100% happy. The total number of stimuli was 22 which comprised the unique 100% emotion crowd stimuli (4) and the stimuli containing different proportions of emotions (18), across all six category pairs.

The experiment was programmed using Presentation™ software. Four response buttons on a standard keyboard, namely "z", "x", ">" and "/" were associated with each of the emotional categories of anger, fear, happy and sad respectively. To help the participant learn these associations we marked each response key with a schematic drawing of a face expressing one of the four emotions. The experiment was run on an Apple Macintosh computer. Stimuli were presented on a Samsung 19″ LCD widescreen display screen with a resolution of 720 dpi. Although head position was not constrained, participants sat such that their head was approximately 57 cm away from the screen. The approximate visual angle subtended by the crowd stimuli as a whole was 16° from central fixation. The image of each character within the crowd subtended a visual angle of approximately 2.5° vertically and 1° horizontally.

Design

The experiment was based on a within subjects design with emotional category (anger, fear, happiness, sadness) and proportion of characters in the crowd conveying one of the emotions in a emotion pair (0, 25, 50, 75, and 100%) as the main factors.

There were two phases to the experiment: a training session followed by the test phase. In the training session, participants learned to associate particular response buttons to each of the four emotional categories until they reached criterion performance. In the subsequent test phase, the experiment was based on a four alternative forced choice paradigm where participants were required to categorise the emotion of each of the crowd stimuli as one of the four category emotions. We measured reaction times and categorisation responses. Each of the 22 crowd stimuli was presented twice over two experimental blocks, i.e one stimulus per block. Trials were randomly presented within each block and across all participants. Participants received a short practice session prior to the test session which consisted of four trials, each displaying stimuli with 100% proportion of each emotion. These practice trials were randomly presented across participants.

Procedure

Prestudy

Prior to the main study, we asked a separate group of 58 participants to rate our crowd stimuli according to the perceived intensity of the emotion portrayed. We conducted this study in order to evaluate whether any effects observed in the main study could be due to the emotional category per se rather than the differences in the intensity of the emotion (i.e. if "happy" crowds were more happy than "angry" crowds were angry). Participants were presented with 100% consistent crowd stimuli, one at a time, and were required to rate each stimulus for its emotional intensity on a scale from 1 to 7. To avoid any bias in the responses to each stimulus, participants rated each crowd

Table 1 Mean ratings provided to each of the emotional crowd stimuli across different categories of emotion in the prestudy of Experiment 1. The highest ratings were given to the one of the four emotional categories used in our study. Also shown are the ratings provided for the second highest and third highest emotional categories related to each of the emotions depicted in the crowd stimuli

Emotion of crowd stimuli	Emotional category of highest ratings	Emotional category of second highest ratings	Emotional category of third highest ratings
Anger	Anger (mean = 6.06, SD = 1.3)	Disappointment (mean = 5.46, SD = 1.66)	Worried (mean = 3.22, SD = 1.66)
Happy	Happy (mean = 6.18, SD = 1.23)	Excited (mean = 6.13, SD = 1.59)	Pleased (mean = 5.70, SD = 1.38)
Fear	Fear (mean = 6.15, SD = 1.46)	Afraid (mean = 5.89, SD = 1.76)	Worried (mean = 4.87, SD = 1.96)
Sad	Sad (mean = 5.91, SD = 1.46)	Disappointment (mean = 5.91, SD = 1.61)	Worried (mean = 4.6, SD = 1.8)

scene along 16 different emotional intensities including anger, afraid, sad, worried, happy, amused, pleased, content, interested, excited, bored, relaxed, disappointed, confident, loving, fear (see Cowie and Cornelius 2003 for a similar procedure). We found that ratings provided to each of the four main emotional categories were higher than those given to other emotional categories that were semantically related to the main emotion (see Table 1). We conducted a one-way, repeated measures ANOVA on the ratings to each of the emotional categories and found a significant effect of Emotion category across ratings [F (3,171) = 8.5, $P < 0.001$]. Tukey post hoc analyses revealed that correct category ratings to the sad crowds were lower than to any of the other emotions ($P < 0.05$), but there was no difference in ratings across the happy, anger and fear crowds. Our ratings here are consistent with previous studies which have also reported that the emotional category 'sad' is typically associated with a low intensity (e.g. Russell 1980).

Main study

The task for participants was to categorise, as fast and as accurately as possible, the overall emotion displayed by the crowd of characters. Participants first learned to associate response keys with each of the emotional categories during a training session. A trial in the training session consisted of the name of an emotional category appearing on a screen and participants were required to press the associated response key as fast and as accurately as possible. This training block was repeated until participants achieved a criterion accuracy response rate of 90%. On average, the block was repeated three times for this criterion to be attained across participants.

Following a short practice session, participants were then presented with the test phase. In this phase, a trial began with a central fixation cross which was presented for 1,500 ms on which participants were required to focus. A crowd stimulus was then presented and participants were required to categorise the stimulus as angry, fearful, happy or sad. Each stimulus was 2 s in duration but this was repeated a maximum of three times during a trial. Otherwise a response triggered the offset of the trial. Participants were instructed to base their decision on the overall emotion depicted by the characters in the scene. There were two blocks to the test phase with 22 trials in each block. Participants could take a (self-timed) break between blocks. The entire experiment took about 15 min for each participant to complete.

Results

The average accuracy at categorising the emotions was 92% across participants and performance was therefore well above chance level of 25%. The average reaction time across all participants was 4,079 ms. Trials in which an incorrect response key was pressed (i.e. not one specified as a response key) were removed from the analysis. This resulted in a removal of 0.7% of the data across all trials. Response times were analysed for outliers (anything above or below 3 interquartile ranges of the mean) leading to the removal of a further 2.4% of the reaction time data.

We first analysed the accuracy and response time performance across each of the crowd stimuli depicting a consistent (i.e. 100%) emotion. The mean accuracy performance for each of the emotional categories was as follows: anger (mean = 90.8%, SD = 2.05), fear (mean = 89.1%, SD = 2.99), happy (mean = 96.67%, SD = 1.76); and sad (mean = 92.22%, SD = 2.66). A one-way, repeated measures ANOVA was performed on these accuracy data with category of emotion as the main factor (anger, fear, sad, happy) and revealed a main effect [F (3,651) = 3.58, $P = 0.01$]. Tukey post hoc analyses found that performance was significantly more accurate for happy crowds than for either angry ($P < 0.05$) or fearful crowds ($P < 0.05$).

Response times to the correct trials only (i.e. hits) were as follows (in milliseconds) for each of the emotions: anger (mean = 3,859, SD = 1,669), fear (mean = 3,307, SD = 1,556), happiness (mean = 3,270, SD = 1,753), sadness (mean = 3,244, SD = 1,559). A main effect of emotional category was found [F (3,588) = 14.01, $P < 0.001$]. Tukey post hoc analyses revealed that participants were significantly slower at categorising anger than

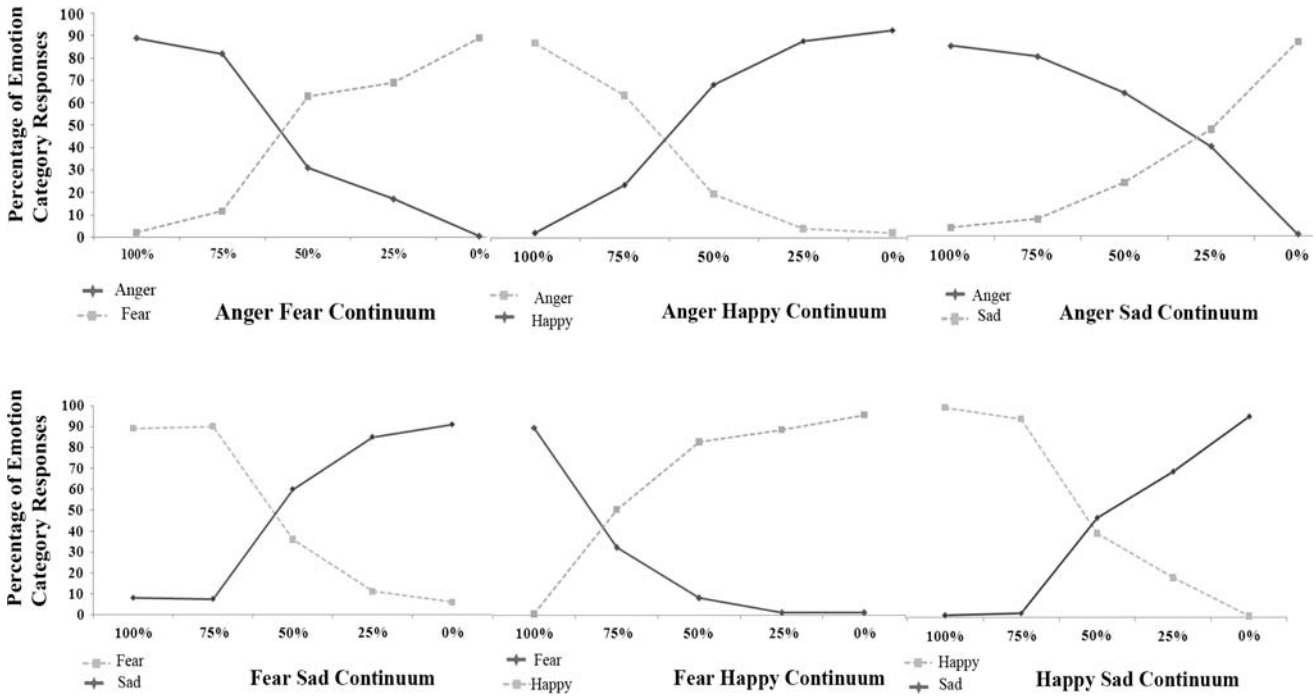

Fig. 2 Mean percentage of categorisation responses for emotions in each of the six emotional continua in Experiment 1

either fear ($P < 0.001$), happiness ($P < 0.001$) or sadness ($P < 0.001$).

Figure 2 shows the percentage of categorisation responses made to each emotional category within each of the six continua. The first observation from these data is that none of the emotions in the crowd appear to be perceived categorically, according to the hallmarks of categorical perception (see, e.g. Young et al. 1997). If emotions of crowds were perceived as categorical we would expect that each emotion would be categorised consistently across participants 100% of the time, even for those emotional pairs in which only 75% of the characters displayed a certain emotion. Furthermore, categorisation responses should change abruptly from one emotional category to another along the continuum, e.g. at the 50% consistent emotion. These effects do not appear in our data but, rather, some emotions seem to dominate the perception of the emotion in a crowd, even when the emotion of the crowd is ambiguous (i.e. 50% consistent). For example, when crowds are conveying 50% of one emotion in a scene, categorisation responses are often higher for one emotion over the other. This is the case for the fear–anger, happy–anger, happy–fear, sad–fear and anger–sad continua where "fear", "happy", "happy", "sad" and "anger" responses dominate respectively. Only in the happy–sad continuum do the categorisation responses reflect the proportion of emotion displayed.

Finally, we compared categorisation data and response time performance across the different proportion of emotions in the crowds. All crowd sizes were collated into three categories for this analysis; 100% one emotion/0% the other emotion, 25% one emotion/75% the other emotion, and 50% one emotion/50% the other emotion. To analyse categorisation responses, we arbitrarily assigned one of the emotions in each crowd as emotion 'A' and the other as emotion 'B' and classified responses in terms of percentage emotion 'A' responses. See Fig. 3 for a plot of the mean performance.

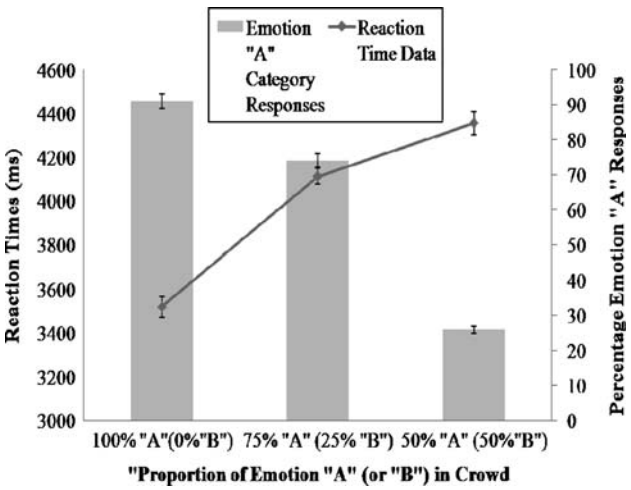

Fig. 3 Plot showing the mean response times and mean Emotion "A" categorisation rates collapsed across each of the proportions of emotions displayed in a crowd in Experiment 1. Here, "A" was arbitrarily assigned to one of the two emotions in the crowd (the *error bars* represent ±1 SE of the mean)

We conducted separate linear regression analyses on the effect of the proportion of emotion in the crowd on both reaction times and categorisation data. Significant effects of the proportion of emotion on reaction times [$R = 0.386$, $F(1,676) = 113.34$, $P < 0.001$] and on the percentage of Emotion "A" responses were found [$R = 0.762$, $F(1, 676) = 933.9$, $P < 0.001$].

Discussion

Participants are highly accurate at identifying the emotion from a crowd when all characters are conveying the same emotion. This suggests that body motion and posture are efficient cues for conveying emotions in the absence of facial expressions. Furthermore, the dynamic cues were presented using virtual humans, in which other spatial cues were controlled, suggesting that virtual characters are efficient portrayers of emotional body language.

In accordance with the literature on the perception of emotional body language in individuals, we found that some emotions were more efficiently perceived from crowds than others. For example, the 'happy' emotion was the fastest and most accurate to categorise, followed by the 'sad' emotion. Both 'anger' and 'fear' had similar categorisation accuracy although 'anger' was slightly slower to categorise. Given that 'sad' was categorised relatively efficiently, it is interesting to note that this emotion was rated as the least intense in our prestudy. These results therefore suggest that it was the emotional category per se, rather than the degree of emotional intensity, which affected responses.

The relative proportion of emotions conveyed by the characters also affected the speed and accuracy at which the overall emotion was perceived. In effect, the greater the proportion of characters displaying a particular emotion the faster and more accurate the categorisation responses of that emotion. Our results therefore suggest that there is a monotonic increase in reaction times and decrease in categorisation of emotion 'A' with a decrease in the proportion of characters in the crowd depicting an emotion. This finding is not consistent with categorical perception which would, in contrast, predict an abrupt shift in responses at some point along the continuum, rather than a gradual shift in categorisation and response times. Furthermore the fact that a difference was found in response times between the 100% and 75%/25% condition, both of which contain a majority of one emotion, is not characteristic of categorical perception, which would predict a difficulty in the ability to differentiate between points within a category. The categorisation responses within each continuum of emotion pairs are also inconsistent with the idea that emotion of a crowd is categorically perceived. Instead, the proportion of categorisation responses to the emotionally ambiguous crowds (i.e. those with 50% of characters depicting one or the other emotion in the continuum) suggests that some emotions dominate the perception of the emotion of the crowd over others. For example, in the anger–sad continuum, categorisation responses changed from 'anger' to 'sad' only when the crowd contained 75% or more characters displaying sadness. Similarly in the fear–happy continuum, the emotion of the crowd was categorised more as 'happy' than 'fearful' when only 25% of the characters conveyed happiness. It appears that happy and angry emotional body expressions may be more salient or dominant in crowd scenes than fear and sadness. Since sad crowds were rated as less intense than the other emotional crowds, we were concerned that other emotions would dominate sadness in the paired continua of emotional crowd scenes. However, apart from anger dominating responses over sad in the anger–sad continuum, it was not generally the case that sad was dominated by other emotions. Thus our findings suggest that some emotions dominate the perception of the emotion of a crowd and these emotions are unrelated to the intensity of the emotion portrayed.

Because of the relatively long exposure to each of the crowd stimuli, it is not clear from the data described here whether or not the emotion of a crowd can be perceived from a brief presentation of the stimulus. In the present study, the stimulus was repeated three times, for a maximum of 6 s and we found that participants took, on average, over 4 s to respond (i.e. after two presentations of the stimulus). To assess the effect of stimulus duration, we decided to reduce the exposure of the crowd stimuli to one repetition, i.e. 2 s in a follow-on study of Experiment 1. A naïve group of 24 participants (13 female) took part in this follow-on study. Apart from the stimulus duration times, the design and procedure used here were identical to that of the previous study. We found that while the mean reaction time fell from 4,079 to 2,403 ms, response patterns were highly similar to those found in the first study.[1] Categorisation accuracy of the 100% crowds was on average 92.7% which was comparable to that found in Experiment 1. Furthermore, the frequency of 'happy' responses was greater than other responses (i.e. 32% of all responses) with lower frequency of 'anger' (25%), 'fear' (20%) and 'sad' (21%) responses. Thus, the patterns of emotional categorisation are very similar across Experiment 1 and the follow-on study, with happiness tending to influence responses to a bigger extent than any other emotion category.

[1] Unequal sample sizes prevented us from conducting statistical comparisons across the two studies.

Overall, these findings suggest that the emotion of a crowd is rapidly detected and that the initial perception of this emotion does not seem to change with longer exposure to the crowd stimulus. This would be in accordance with research on the perception of emotional gist which has suggested that the judgement of the emotion of a scene can be arrived at in as little as 200 ms (Righart and de Gelder 2008).

Experiment 2

In the following experiment, we investigated whether the emotion of a crowd was affected by the number of characters in the crowd, and not just the proportion of characters conveying the emotion. We expected that response times in particular would increase with larger display sizes if a character-by-character search was conducted in order to determine the overall emotion of the crowd. Otherwise, if the emotion of a crowd is determined without the necessity of monitoring each of the individual characters in the display, then response times should not vary as a function of crowd size.

Method

Participants

We again recruited an opportune sample of 70 participants (38 female) from attendees at an exhibition at the Trinity College Science Gallery, Dublin. None of these participants took part in the previous experiment. Their ages ranged from 16 to 67 years, with a mean age of 32.19. All participants gave informed written consent to partake in the study, and all reported normal or corrected-to-normal vision.

Stimuli

We used three different crowd size displays as stimuli in our experiment; one with 12 characters, another with 20 characters (as in the previous experiments) and one with 32 characters. We maintained the same stimulus proportions and visual angles as described in Experiment 1.

All crowd stimuli were created as described in Experiment 1. Our stimulus set resulted in 66 crowds. Each stimulus was presented for a duration of 2 s.

Design and procedure

This experiment was based on a within subjects design with emotion category (anger, fear, happy, sad), proportion of characters displaying the emotion (0, 25, 50, 75, 100%) and crowd size (12, 20 or 32 characters) as the main factors. As in the previous experiments we measured reaction times and categorisation responses. The procedure was the same as described in Experiment 1.

Results

Missing values accounted for 1% of all responses. On average, 87% of the emotions of crowds were correctly categorised in the 100% consistent crowds stimuli only. Outliers were first removed from the response times prior to any analysis (using the same criteria as in Experiment 1), which led to a removal of 1% of all response time data. The mean overall response time taken to respond to these consistent crowd stimuli was 2,229 ms.

Our results for the categorisation responses across each of the six continua were similar to those found in Experiment 1. For example, responses to the 50% ambiguous crowds in the anger–fear, anger–happy, fear–happy, anger–sad, fear–sad and happy–sad continua followed similar patterns across experiments with some emotions dominating the perception of the emotion of the crowd. The only differences across experiments were in the fear–happy continuum in that here happy dominated the perception of the emotion of the crowd when 25% of the characters were happy instead of 50% as in Experiment 1, and in the fear–sad continuum, fear dominated the perception when 50% of the characters were fearful, instead of only 25% in the previous experiment. It is possible that these differences were due to the different crowd sizes across experiments.

To assess whether responses were categorical across the emotional continua, we pooled the categorisation responses by arbitrarily assigning one of the emotional categories in a paired continuum as 'emotion A' and the other as emotion 'B'. Figure 4 shows the mean accuracy and response time

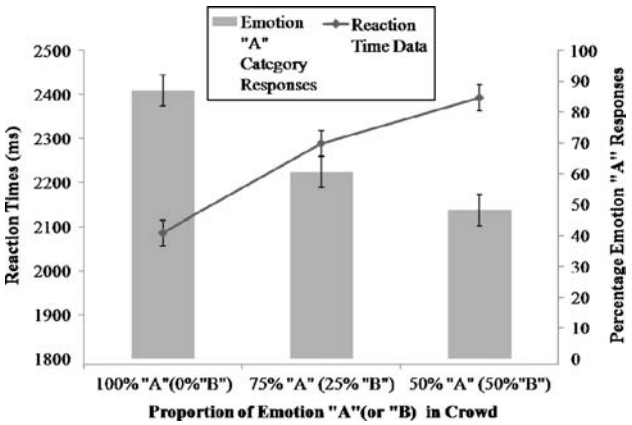

Fig. 4 Plot showing the mean response times and mean Emotion "A" categorisation rates across each of the proportions of emotions displayed in a crowd in Experiment 2 (the *error bars* represent ±1 SE of the mean)

performance across the different proportions of emotions in the crowds. We then conducted a regression analyses on the percentage of Emotion "A" categorisation responses and found a significant effect [$R = 0.149$, $F(1, 207) = 4.7$, $P < 0.05$]. A similar analysis was conducted on the response time data which was also significant [$R = 0.428$, $F(1, 207) = 46.06$, $P < 0.001$]. Thus we found increasing response times and decreasing number of emotion "A" category responses with a decrease in the ratio between emotion "A" and "B" across all the continua.

We then assessed the effect of crowd size and the proportion of characters displaying a consistent emotion on perceptual performance by analysing the response times and categorisation data using separate 3 × 3 repeated measures ANOVAs with crowd size (12, 20 and 32 characters) and the overall proportion of characters displaying an emotion (100, 75 or 25 and 50%) as factors. For the categorisation data (i.e. the percentage of emotion "A" categorisation responses), main effects of crowd size [F (2, 136) = 16.13, $P < 0.001$] and proportion of characters [F (2, 136) = 93.87, $P < 0.001$] and a significant interaction between these factors [F (4, 272) = 20.73, $P < 0.001$] were found. Post hoc Tukey analyses on the main effects revealed more categorisation of emotion "A" responses to the 100%/0% crowd proportions than either the 75%/25% or the 50%/50% crowd ($P < 0.001$) and more categorisation "A" responses to the medium crowd sizes than to the small or large crowd sizes ($P < 0.001$). A post hoc analysis of the interaction revealed more emotion "A" categorisation responses to the medium crowd size than either the large or small crowd sizes for all crowd proportions except the 100% consistent crowd where there were more responses to the medium than the large crowds only.

For the response times data, a significant effect of proportion of emotion was found [F (2, 132) = 849.0, $P < 0.001$] with faster response times to the 100% crowds than to either the 25%/75% ($P < 0.001$) or the 50%/50% proportion condition ($P < 0.001$). We also found an effect of crowd size [F (2,132) = 28.8, $P < 0.001$] with faster response times to the small (12) than large crowd (32) and slower response times to the medium (20) than either the small or large sized crowd ($P < 0.05$). A significant interaction between the two was also found [F (4, 2216) = 2.66, $P = 0.03$] (see Fig. 5). For both the 75%/25% crowds and the ambiguous 50% crowds, response times to the medium sized crowd were faster than to either the small or large crowd ($P < 0.001$). For the 100% consistent crowds, response times were faster to the small crowds than either the medium or large crowd sizes ($P < 0.001$). In sum, although response times were fastest to the small crowds and slowest to the medium size crowds, the categorisation responses differed across crowd sizes with more of a bias in responses towards one particular emotion in the medium size crowds of 20 individuals than the other two crowd sizes.

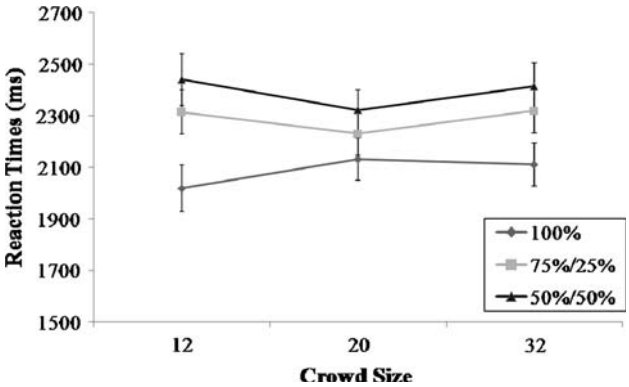

Fig. 5 Plot showing the mean response times as a function of crowd sizes across each of the proportion of emotions conveyed by the crowd in Experiment 2 (the *error bars* represent ±1 SE of the mean)

Finally, we analysed the accuracy performance across the 100% consistent emotion crowds (anger, fear, happy and sad) and the crowd size (12, 20 or 32 characters) using a 4 × 3 repeated measures ANOVA. Significant main effects of crowd size [F (2, 138) = 5.75, $P < 0.01$] and emotional category [F (3, 207) = 7.7, $P < 0.001$] were found. Performance to the happy crowds was significantly more accurate than to either the anger (post hoc Tukey test, $P < 0.05$) or fear ($P < 0.001$) crowds and performance to sad crowds were more accurate than to fear crowds ($P < 0.005$). Furthermore, performance was least accurate to the crowds containing 20 characters than to the crowds containing either 12 or 32 characters ($P < 0.05$). A significant interaction between the two factors was also observed [F (6, 414) = 4.41, $P < 0.001$] (see Fig. 6a). Post hoc tests on the interaction revealed that for the 'anger' crowds, performance was more accurate for the 12 and 32 character crowds than the 20 character crowds ($P < 0.01$). Performance for the 'fear' crowds was also similar in that performance was better to the 12 characters than 20 character crowds ($P < 0.001$) but no different to 32 characters. There was no difference across crowd sizes for either of the 'happy' or 'sad' crowds.

A similar analysis on the response times revealed main effects of emotion [F (3, 195) = 583.8, $P < 0.001$] and crowd size [F (2, 130) = 133.8, $P < 0.001$]. The time to categorise 'anger' was longer than the time to categorise the other three emotions ($P < 0.001$) and response times to categorise 'fear' were longer than to either sad or happy emotions ($P < 0.001$). The time taken to categorise the emotion of a crowd size of 20 characters was longer than either of the other two crowd sizes ($P < 0.01$) but response times to the smallest crowd (12) were faster than to the largest crowd size (32) ($P < 0.001$). There was also a significant interaction

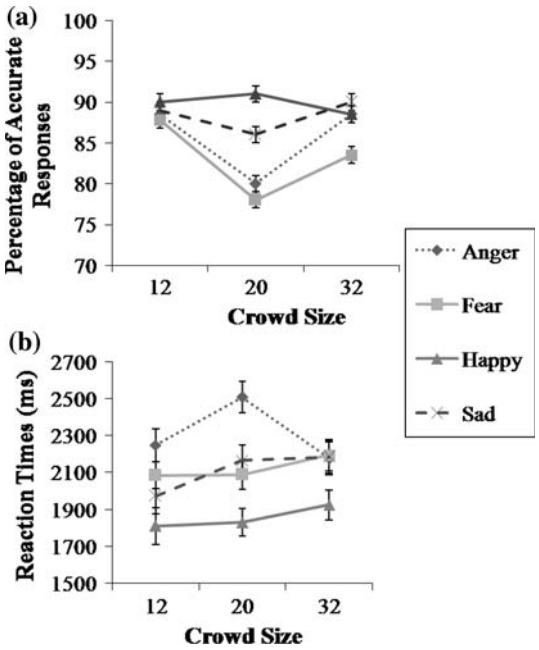

Fig. 6 Plot showing **a** mean accuracy performance and **b** mean response times across the different crowd sizes when the characters in the crowds conveyed the same emotion (i.e. 100% emotion crowds) in Experiment 2 (the *error bars* represent ±1 SE of the mean)

between these factors [F (6, 390) = 48.8, $P < 0.001$] (see Fig. 6b). Tukey post hoc analyses revealed that the effect of crowd size differed across emotional categories: for the 'fear' emotion, response times were faster to the small than large crowd sizes ($P < 0.01$); for the 'anger' emotion response times were slowest to the medium than either the small or large crowds ($P < 0.01$); for the 'happy' emotion, response times were slowest to the large crowds than either the small or medium crowds ($P < 0.01$) and for the 'sad' emotion response times were faster for the small than the medium or the large crowds ($P < 0.01$).

Discussion

The aim of Experiment 2 was to clarify the effect of crowd size on the perception of the emotion of a crowd. As in the previous experiment, the accuracy at categorising the emotion of the consistent crowds was high (i.e. 87%) and the time taken to determine the emotion was faster for the consistent than the mixed crowds. Thus the perception of the emotion of crowds of varying sizes was overall very efficient.

We found that emotion was categorised better and faster for some crowd sizes than others. For example, participants were generally more accurate and faster at categorising the emotion displayed by crowds of 12 characters or 32 characters than by crowds of 20 characters, especially when the emotion displayed in the crowd was either anger or fear. It is not clear why responses to the medium sized crowds were relatively less efficient than other crowd sizes for some emotions than others. One suggestion may be that different crowd sizes or crowd densities may be required for the optimal portrayal of some emotions. Further research would help clarify how crowd size or crowd density affects the perception of different types of emotions. More pertinently, the overall pattern of categorisation responses to each emotional continuum was very similar to those patterns found in Experiments 1 (with the exception of a small shift in responses in the fear–happy continuum), suggesting that some emotions dominate perception irrespective of crowd size.

Our results therefore suggest that the perception of emotion from crowds is not affected in a systematic way by the size of the crowd. In other words, we failed to find evidence to suggest that emotion is more difficult and slower to determine with increasing number of characters in a crowd. Instead, emotion categorisation responses were most efficient to the small and large crowd sizes. These findings suggest that an item-by-item search through each of the characters in the crowd display did not occur and that the emotion conveyed by a crowd was either determined in a parallel fashion or that an optimal number of characters were sampled in each scene irrespective of the overall crowd size.

General discussion

The main aim of the current study was to investigate how the emotion of a crowd is perceived based only on the body expressions of characters in the crowd. We found, in two experiments, that the emotion of a crowd containing one emotion only can be accurately perceived and rapidly determined, even with relatively brief exposure of the crowd scenes. As would be expected, the crowds containing one emotion were also faster to be categorised than those containing two emotions. Moreover, although crowd size had an effect on the categorisation of the emotion in a crowd, this was not because the emotion of increasing crowd sizes was more difficult to categorise. Rather, we found that emotion was easier to categorise from relatively small and large crowds than medium sized crowds.

We also found a monotonic increase in categorisation responses with an increase in the proportion of characters depicting one emotion over another. In other words, our results did not suggest that categorisation responses rapidly shifted from one emotional category to another along the emotional continua. Instead, we found a systematic difference in categorisation responses and response times

across experiments between the 100%, 75%/25% and 50/50% crowd proportions, suggesting strong evidence that the emotions in crowds are perceived according to the consistency with which the emotion is conveyed across the characters rather than as discrete categories. It is possible that more inter-category steps may reveal any emerging categorical effects in our task. However, even with relatively few steps along each continuum, if the emotions of a crowd were perceived as categorical then we would still expect little difference in categorisation performance between proportions of crowds that lay on one side of the category boundary. In the absence of any evidence for such within category equivalent performance, our results suggest that the perception of the emotion of a crowd is not categorical.

Our results here are seemingly at odds with those of Young et al. (1997) who found strong evidence for categorical perception of facial expressions. It may be that facial expressions are categorical whilst emotional body language and crowd perception are not. Moreover, there may be important reasons why the emotional state of a crowd is not perceived as categorical. Indeed, it may be more beneficial for survival if the emotion conveyed by even a small number of characters in a scene is directly perceived. As such, the perception of emotion in groups of individuals may rely on different mechanisms than the perception of emotion in a single individual.

Our results also suggest that some emotions dominate the perception of the emotion of a crowd relative to others. For example, we found that the 'happy' emotion was more efficiently categorised than any of the other three emotions expressed in the crowds. Stalans and Wedding (1985) explain the pop-out effect of an angry face in a happy crowd by suggesting that the happy distracters are processed much faster than negative stimuli. This could explain the fast response times to the happy crowds. Furthermore, we noticed that happy characters appeared to move relatively more than other emotions, e.g. the characters jumped and clapped hands, which may have made it more perceptually salient than the other emotions. The fact that happiness as a category judgement appeared to dominate over most other emotions throughout all three studies would also be indicative of differences in perceptual saliency. In contrast, although sad characters moved less and with less exaggerated movements than the other emotive characters there was no difference in categorisation responses between sad and the other emotions of fear and anger. This suggests that the rate of body movements alone was not the only cue used to perceive the emotion of the crowd.

The relative dominance of happy over other emotions in crowds was an unexpected result, since, according to the 'alarm' hypothesis other threat-related emotions are thought to be processed preferentially (Walk and Homan 1984). For example, activations in the amygdala are thought to be related to vigilance for fear-related stimuli in the environment (see, Calder et al. 2001). However, previous work refers to the perception of single entities expressing emotion. It could be that the neural mechanisms involved in determining the emotion from a crowd is different to those involved in the perception of emotion in individuals. Clearly, future neuroimaging studies comparing the perception of emotion from individuals or groups of individuals could help elucidate the neural mechanisms underpinning each percept.

In Experiment 1, we reduced the duration of the crowd stimuli from 6 to 2 s and found that although response times were consequently reduced, there was also a decrease in how accurate the emotions were categorised in the 100% consistent crowds only. These results suggest that the perception of emotion in crowds relies on the accurate analysis of the emotional expression of some of the characters contained within the crowds. However, when we manipulated crowd size in Experiment 2, we found no systematic relationship between the number of characters in the crowd and categorisation accuracy, suggesting that a small proportion of the characters in a crowd may be attended to in order to derive the general emotion of the crowd. Furthermore, we found that patterns of categorisation responses were generally unaffected by either a change in duration or in crowd size across experiments. However, it is unclear from our data whether brief exposure to the crowd scene would further decrease performance. For example, Righart and de Gelder (2008) report that a target emotion is detected in a scene with just 200 ms presentation duration. Thus it remains possible that the emotion of crowds would remain highly accurate even at shorter durations than those tested here.

We manipulated crowd size and found that participants were more efficient at identifying the emotion of crowds containing 12- or 32 characters than the 20 character crowds. This was an unexpected result since if serial processing were occurring then we would expect response times to increase with increasing crowd sizes. One suggestion is that different mechanisms are effecting the relatively better response times to the small and large crowds. For example, determining the emotion of a small crowd may be relatively easy relative to larger crowd sizes because of the small number of characters that need to be attended. With regards to larger crowd sizes, this increasing density in crowds may benefit perception since characters are spatially closer to each other. Furthermore, the effect of crowd size was strongest for the angry crowds than other emotions, suggesting that some emotions may be more readily determined from crowds irrespective of the number of characters in the crowd.

The results of the current study suggest that the emotional body expression of characters in a crowd is an effective cue for determining the overall emotion of the crowd. The emotion expressed by varying proportions of the characters in a crowd had a direct effect on the overall emotion perceived. That is, when crowds contained characters conveying different emotions, we found that these different emotions were detected in the scene. Furthermore, although some emotions dominated the perception of the emotion of a crowd, this did not lead to the categorical perception of emotion. The perception of emotion in crowds, therefore, seems to rely on a different mechanism to the perception of emotion in individuals based on either facial expression or body language. Our study provides an initial investigation of an important social cue, namely how we perceive the emotion across a group of individuals, and our findings have important implications for our understanding of how we recognise and act upon emotional crowds.

References

Adolphs R (2002) Neural systems for recognizing emotion. Curr Opin Neurobiol 12(2):169–177

Allison T, Puce A, McCarthy G (2000) Social perception from visual cues: role of the STS region. Trends Cogn Sci 4(7):267–278

Calder AJ, Lawrence AD, Young AW (2001) Neuropsychology of fear and loathing. Nat Rev Neurosci 2:352–353

Cowie R, Cornelius R (2003) Describing the emotional states that are expressed in speech. Speech Commun 40(1–2):5–32

De Gelder B (2006) Towards the neurobiology of emotional body language. Nat Rev Neurosci 7(3):242–249

De Gelder B, Hadjikhani N (2006) Non-conscious recognition of emotional body language. Neuroreport 17(6):583–586

De Gelder B, Vroomen J, Pourtois G, Weiskrantz L (1999) Non-conscious recognition of affect in the absence of striate cortex. Neuroreport 10(18):3759–3763

De Gelder B, Morris JS, Dolan RJ (2005) Unconscious fear influences emotional awareness of faces and voices. Proc Natl Acad Sci USA 102:18682–18687

Ekman P (1973) Darwin and facial expression: a century in review. Academic Press, London

Ekman P (1982) Felt, false and miserable smiles. J Nonverbal Behav 6(4):238–252

Ekman P (1992) An argument for basic emotions. Cogn Emot 6:169–200

Etcoff NL, Magee JJ (1992) Categorical perception of facial expressions. Cognition 44(3):227–240

Fabri M, Moore D, Hobbs D (2004) Mediating the expression of emotion in educational collaborative virtual environments: an experimental study. Virtual Real 7(2):66–81

Hampton C, Purcell DG, Bersine L, Hansen CH, Hansen RD (1989) Probing "Pop-out": another look at the face-in-the-crowd effect. Bull Psychon Soc 27(6):563–566

Hansen CH, Hansen RD (1988) Finding the face in the crowd: an anger superiority effect. J Pers Soc Psychol 54:917–924

Harnad S (ed) (1987) Categorical perception: the groundwork of cognition. Cambridge University Press, New York

McDonnell R, Joerg S, McHugh J, Newell F, O'Sullivan C (2009) Investigating the role of body shape on the perception of emotion. ACM Trans Appl Percept (in press)

Morris JS, de Gelder B, Weiskrantz L, Dolan RJ (2001) Differential extrageniculostriate and amygdala responses to presentation of emotional faces in a cortically blind field. Brain 124(6):1241–1252

Newell FN, Bülthoff HH (2002) Categorical perception of familiar objects. Cognition 85(2):113–143

Reed CL, Stone VE, Bozova S, Tanaka J (2003) The body inversion effect. Psychol Sci 14(4):302–308

Righart R, de Gelder B (2008) Facial expressions and emotional scene gist. Cogn Affect Behav Neurosci 8(3):264–272

Russell JA (1980) A circumplex model of affect. J Pers Soc Psychol 88:3–32

Scherer KR (1981) Speech and emotional states. In: Darby J (ed) Speech evaluation in psychiatry. Grune & Stratton, New York, pp 189–220

Schindler K, Van Gool L, de Gelder B (2008) Recognising emotions expressed by body pose: a biologically inspired neural model. Neural Netw 21(9):1238–1246

Stalans L, Wedding D (1985) Superiority of the left hemisphere in the recognition of emotional faces. Int J Neurosci 25(3–4):219–233

Stekelenburg JJ, de Gelder B (2004) The neural correlates of perceiving human bodies: an ERP study on the body-inversion effect. Neuroreport 15(5):777–780

Sturman DJ (1994) A brief history of motion capture for computer character animation. Course Notes 9, Character Motion Systems, ACM SIGGRAPH 1994, Los Angeles, USA

Treisman A, Gelade G (1980) A feature-integration theory of perception. Cogn Psychol 12(1):97–136

Walk RD, Homan CP (1984) Emotion and dance in dynamic light displays. Bull Psychon Soc 22(5):437–440

Walk RD, Walters KL (1988) Perception of the smile and other emotions of the body and face at different distances. Bull Psychon Soc 26(6):510

Wallbott H (1998) Bodily expression of emotion. Eur J Soc Psychol 28:879–896

Whalen PJ, Rauch SL, Etcoff NL, McInerney SC, Lee MB, Jenike MA (1998) Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. J Neurosci 18(1):411–418

Young AW, Rowland D, Calder AJ, Etcoff NL, Seth A, Perrett DI (1997) Facial expression megamix: tests of dimensional and category accounts of emotion recognition. Cognition 63(3):271–313