Dynamic facial expressions are processed holistically, but not more holistically than static facial expressions

Alanna Tobin
University of Wollongong, atobin@uow.edu.au

Simone K. Favelle
University of Wollongong, skeane@uow.edu.au

Romina Palermo
University of Western Australia

Dynamic facial expressions are processed holistically, but not more holistically than static facial expressions

Abstract

There is evidence that facial expressions are perceived holistically and featurally. The composite task is a direct measure of holistic processing (although the absence of a composite effect implies the use of other types of processing). Most composite task studies have used static images, despite the fact that movement is an important aspect of facial expressions and there is some evidence that movement may facilitate recognition. We created static and dynamic composites, in which emotions were reliably identified from each half of the face. The magnitude of the composite effect was similar for static and dynamic expressions identified from the top half (anger, sadness and surprise) but was reduced in dynamic as compared to static expressions identified from the bottom half (fear, disgust and joy). Thus, any advantage in recognising dynamic over static expressions is not likely to stem from enhanced holistic processing, rather motion may emphasise or disambiguate diagnostic featural information.

Keywords

than, more, not, but, static, holistically, dynamic, processed, expressions, facial

Disciplines

Education | Social and Behavioral Sciences

Publication Details

Dynamic facial expressions are processed holistically, but not more holistically than static facial expressions.

Alanna Tobin[1*], Simone Favelle[1] and Romina Palermo[2]

[1]School of Psychology, University of Wollongong, Australia;

[2]ARC Centre of Excellence in Cognition and its Disorders (CCD) and School of Psychology, University of Western Australia, Australia

RUNNING HEAD: Holistic processing of dynamic emotions

*Correspondence to:

Alanna Tobin

School of Psychology

University of Wollongong

Northfields Ave, Wollongong 2522

NSW, Australia

Telephone: +61 2 4221 3742

Email: atobin@uow.edu.au

Abstract

There is evidence that facial expressions are perceived holistically and featurally. The composite task is a direct measure of holistic processing (although the absence of a composite effect implies the use of other types of processing). Most composite task studies have used static images, despite the fact that movement is an important aspect of facial expressions and there is some evidence that movement may facilitate recognition. We created static and dynamic composites, in which emotions were reliably identified from each half of the face. The magnitude of the composite effect was similar for static and dynamic expressions identified from the top half (anger, sadness, surprise) but was reduced in dynamic as compared to static expressions identified from the bottom half (fear, disgust, joy). Thus, any advantage in recognising dynamic over static expressions is not likely to stem from enhanced holistic processing, rather motion may emphasise or disambiguate diagnostic featural information

Keywords: emotion recognition, facial motion, composite effect, holistic, features, inversion effect

Recognising the emotion expressed on other people's faces is important for social and interpersonal interactions, is typically accomplished easily, and deficits are associated with psychiatric and neurological disorders (e.g., Kennedy & Adolphs, 2012). It is often argued that there are six "basic" expressions of emotion - joy, fear, disgust, sadness, surprise and anger - that are recognised cross-culturally (Ekman & Friesen 1978, but see Feldman Barrett, 2006). Here, we seek to understand whether the perceptual processes involved in the recognition of these six expressions differ when they are static as compared to moving.

The perception of emotion is undeniably complex (see Beaudry, Roy-Charland, Perron, Cormier, & Tapp, 2014 for recent further discussion), with evidence that facial expression may be coded relative to a norm (Burton, Jeffery, Calder, & Rhodes, 2015; Skinner & Benton, 2010) and may be "embodied" involving both visual processing and simulation of somatosensory brain regions specific to each emotion (Pitcher, Garrido, Walsh, & Duchaine, 2008). In regards to the information that may be used to form visual representations, there is evidence for individual features being necessary and sufficient for the recognition of some expressions, such as an upturned mouth in the recognition of happiness (e.g., Beaudry et al., 2014), whereas other expressions seem to rely upon the processing of multiple features, such as the nose and mouth for disgust (which could be processed sequentially, see Beaudry et al., 2014 or simultaneously but not holistically, see Ellison & Massaro, 1997). There is also evidence that some expressions, such as fear (e.g., Beaudry et al. 2014) may be processed more holistically, in which there is very strong perceptual integration across the whole face simultaneously so that the whole perceived is more than the sum of the parts (see detailed definitions in McKone et al, 2013; Rossion, 2013).

However, an important caveat of this work is that studies typically use static images of facial expressions. However, facial expressions are dynamic events with characteristic

facial movements. Dynamic faces provide rich information not available in static images including motion and temporal information and cues to three-dimensional shape. In fact, the dynamic, changeable nature of facial expressions is why current functional and anatomical models of face recognition separate the processing of expression from that of identity (e.g., Bruce & Young, 1986; Haxby, Hoffman, & Gobbini, 2000). Despite this, findings are equivocal with regard to any dynamic advantage for expression recognition: some provide support for a dynamic advantage (e.g., Cunningham & Wallraven, 2009); some suggest that the advantage may only occur in sub-optimal conditions, such as poor image quality (e.g., Bassili, 1979) or subtle expressions (e.g., Ambadar, Schooler & Cohn, 2005); whereas others find no advantage at all (e.g., Fiorentini & Viviani, 2011).

Here we address whether there are functional differences in processing static and dynamic facial expressions by examining the degree of holistic processing of both types of faces. The findings of this study can contribute to our understanding of facial expression processing in multiple ways. First, we can determine the ecological validity of existing research that has used static stimuli. Second, we can examine whether motion improves recognition by facilitating holistic processing. We may find that the magnitude of holistic processing is equivalent, regardless of whether facial expressions are static or dynamic, indicating that motion confers no holistic processing advantage. On the other hand, because moving faces may exploit Gestalt grouping principles not available in static images, such as synchrony (Piepers & Robbins, 2012), holistic processing may be more evident for dynamic than static facial expressions (or at least for some expressions). Alternatively, holistic processing could be stronger for some/all static as compared to dynamic facial expressions, suggesting that holistic processing may not occur in real-word dynamic interactions. These outcomes should not be taken to indicate that other types of processing, such as featural or embodied processing, are not also occurring. Finally, holistic processing could be absent,

which would be consistent with visual processing proceeding primarily via the use of other strategies, such as feature processing.

One method of investigating holistic processing is to compare recognition of upright and inverted faces, since only upright faces are processed holistically (McKone, et al 2013; Rossion, 2013). Ambadar et al. (2005) found that inversion significantly disrupted recognition of subtle expressions (anger, disgust, fear, joy, sadness, surprise) to the same degree in both static and dynamic stimuli, suggesting that expressions are holistically processed in both. However, inversion is considered only an indirect marker of holistic processing (McKone, Martini, & Nakayama, 2001). A more direct method of assessing holistic processing is the composite task, which involves combining two different facial expressions of the same person (one shown in the top half of the face, one in the bottom) so that when aligned the two halves create a perceptually novel "composite" expression (see Figure 1, left). Typically, participants take longer to identify the expression in the top (or bottom) half of the composite compared to a condition in which the halves are misaligned and not face-like (Figure 1, right). This composite effect is due to the difficulty in attending to half of a face when holistically processing the whole composite.

<center>INSERT FIGURE 1 HERE</center>

Studies using the composite task provide evidence for holistic processing of *static* facial expressions (e.g., Calder, Young, Keane, & Dean, 2000; Palermo et al., 2011; Tanaka, Kaiser, Butler, & Le Grand, 2012). To our knowledge, only one study has investigated holistic processing of dynamic and static expressions using a version of the composite task (Chiller-Glaus, Schwaninger, Hofer, Kleiner, & Knappmeyer 2011). Their study used composites of every combination of the six "basic" expressions (anger, disgust, fear, joy, sadness, surprise), in the top and bottom halves of the face. Analyses of the accuracy of

expression judgements were complex, addressing congruency of expressions in the halves, motion, emotion, and alignment, with no specific comparison of alignment differences for individual static and dynamic expressions. Their results appear to show that some (joy, sadness, surprise in both halves and fear when identified in the top half of the face) but not other (disgust, anger and fear in the bottom half of the face) expressions were holistically processed in both static and dynamic faces. They concluded that holistic processing is not enhanced by motion. However, some limitations of this study warrant a further investigation of this important question. First, it is not clear whether there were significant composite effects in reaction time (RT) data as only analyses of accuracy data were presented (except for joy where accuracy was at ceiling). The composite effect may be found in *either* RT or accuracy and it is more often found in analyses that consider RT, particularly when presentation duration is long (Rossion, 2013). Second, and most importantly, when conducting an expression composite task a naming response is required so it is critical that the expressions are identifiable from the target half of the face (Calder et al., 2000). The half-face expression stimuli used by Chiller-Glaus et al. (2011) had no apparent labeling data and as such, no indication of whether the expressions shown in each half were reliably identifiable. While all above chance, recognition accuracy for some of the expression judgments were quite poor (e.g., anger, fear and disgust in the bottom half were labeled correctly less than 50% of the time) which makes interpretation of their results difficult. Since their version of the composite task combined every top half expression with every bottom half expression, it is possible that some effects may have been obscured.

The primary aim of the current study is to determine whether the composite effect marker of holistic processing, commonly found in static facial expressions, is also found for dynamic expressions. We employed a composite task design that used only combinations of expression halves known to elicit accurate labeling responses (Calder, et al, 2000; Palermo et

al., 2011; Tanaka, et al, 2012). Given that different expressions engage different regions and features to varying degrees, a second aim is to determine whether holistic processing is consistent across expressions (for the face halves selected for this task). In addition we tested inverted composite expressions so that low-level image factors and secondary cognitive factors may be accounted for when interpreting results (McKone et al., 2013). With dynamic faces in particular, it is important to rule out the contribution of the motion signal itself since the motion signal is equally available in upright and inverted dynamic faces. If evidence is found for holistic processing in upright dynamic faces, it is important to be able to determine that it is a face specific process, rather than a cue inherent in the motion itself. We expect to find the composite effect in both static and dynamic expressions. It is less clear whether this will be consistent across expressions.

Method

Participants

30 undergraduate psychology students (7 male) aged from 17 to 44 years ($M = 23.7$) from the University of Wollongong participated in exchange for course credit. Data from an additional 5 participants were excluded due to having more than five cells with zero accuracy ($n = 2$) or having two or more RT cells greater than three standard deviations away from the mean in the upright condition (collapsed across expression, $n = 3$).

Materials

Stimuli were prepared from four models (2 male), from the Amsterdam Dynamic Facial Expression Set (ADFES; van der Schalk, Hawk, Fischer, & Doosje, 2011), each displaying one example of the expressions anger, disgust, fear, joy, sadness, and surprise. Four models were used in line with previous studies (Calder et al., 2000; Palermo et al.,

2011) and to ensure a feasible number of trials given the number of variables manipulated within-subjects (motion, orientation, face-half, alignment). Dynamic composites were 6 s video sequences, starting with the face in a neutral pose with expression onset after an average of 943 ms (SD = 67.4), followed by expression apex after 251 to 695.5 ms (M = 488.4, SD = 63.7). Posers maintained this apex expression with minimal additional movement. Static stimuli were created from a frame of the video sequence at expression apex. Although expression onset varies with the emotion displayed, it is important to note that we compare aligned with misaligned composites, with the same timings in each condition.

Because accurate recognition of expressions in the face halves alone is critical for the composite expression task (Calder et al., 2000), stimuli were selected based on high recognition accuracy for ADFES half-face expressions from a previous study (Favelle, Tobin, DeMayo, & Palermo, 2012) and confirmed in the current experiment (see Appendix A, Table A1). Composites were created by combining each best-recognised top half expression (anger, sadness, surprise) with each best-recognised bottom half expression (disgust and fear; see Figure 1.3). Joy was recognised equally well in each half, and so was included as a bottom-half expression for equal numbers of each half. There were nine combinations of expressions for each of the four models (anger-disgust, anger-fear, anger-joy, sadness-disgust, sadness-fear, sadness-joy, surprise-disgust, surprise-fear, surprise-joy). Composites were prepared using both static (36) and dynamic (36) stimuli. Further, both aligned and misaligned versions of the composites were created for a total of 104 composites.

Misaligned composite stimuli were the same as aligned composite stimuli except that the bottom halves were horizontally offset by half a face width to the left (for half of the stimuli) or the right. Misaligned composites were presented in the centre of the screen (i.e., so that neither the top nor bottom half was presented in the centre of the screen). To ensure

consistency in positioning across conditions, half of the aligned composite stimuli were presented in the same position as the left section of the misaligned composites and half in the same position as the right section (ordered randomly). Inverted stimuli were created by rotating each of the 104 aligned and misaligned composites 180 degrees in the picture plane.

The stimuli were presented in full colour on a 48 cm flat-screen monitor with a resolution of 1024 x 768 pixels. The experiment was run on a Macintosh G5 computer with Psyscope experimental software Version X B57 (Cohen, MacWhinney, Flatt, & Provost, 1993; http://psy.ck.sissa.it/) controlling the trial sequence. The average vertical height and horizontal width of the visible face region (including forehead and some neck) were 12cm and 13cm respectively. At an approximate viewing distance of 50cm, this is a vertical visual angle of 13.5^{o} and horizontal visual angle of 14.8^{o}.

Procedure

Prior to the experiment, a practice block was conducted to familiarise participants with the labelling task and to check expression recognition accuracy (c.f. Calder et al., 2000; Palermo et al., 2011). Participants were asked to identify the expression in each of the 96 whole faces (6 emotions, 4 models, 2 stimulus types [static and dynamic], 2 orientations), presented individually and in random order. Participants made a key-press response using the numbers 1-6. A legend at the bottom of the screen displayed the number corresponding to each expression (label order counterbalanced across participants) [1]. A fixation cross (500 ms) was followed by a blank interval (500 ms) after which the face was presented (with the legend below) and remained on screen until the participant responded. A response initiated the next trial after an interval of 1000 ms.

---

[1] A forced choice response is required in this paradigm (six choices for the familiarization phase and three choices for the composite task). While the reliability of forced choice response in emotion classification has been questioned in the past (Russell, 1993), it is less of an issue here since the composite effect is measured as a difference between aligned and misaligned conditions.

Participants then completed the experiment proper. Face half (top and bottom), stimulus type (static and dynamic) and orientation (upright and inverted) were all blocked separately for a total of eight blocks. The blocks were presented in random order for each participant. Trial type (aligned and misaligned) was randomised within each block. Since the design of these blocks was the same, only a detailed description of the static, upright block identifying the bottom half is given.

As an additional expression recognition check, the block began with presentation of the upright, static bottom halves in isolation, showing each of the three expressions (disgust, fear, joy) posed by the four models. Participants identified the expression with a key-press (numbers 1-3 corresponding to *disgust, fear*, and *joy*). Stimulus and legend remained on screen until the participant responded. Label order was counterbalanced across participants[2]. Experimental trials followed. Each of the 36 aligned and 36 misaligned upright, static composite stimuli were presented individually, in random order, for a total of 72 trials per block (for a total of 576 experimental trials). Participants were asked to identify the expression displayed in the bottom half of the face by key-press, as for the halves in isolation. Trial event timings for both the isolated halves, and the composite faces were the same as for the whole face familiarisation task. Participants were instructed to answer as quickly as possible but also to be accurate.

The experiment session lasted approximately 90 minutes including breaks between each of the blocks and a 10 minute mid-way break in which participants were permitted to leave the testing room. Research protocol was approved by the University of Wollongong Human Research Ethics Committee (reference HE12/079).

---

[2] Analysis of the RT data, with label order as a factor, confirmed that label order did not systematically affect performance as a main effect or interaction (all $F<4.02$, all $p>.05$).

Results

Accuracy for isolated face halves

Expression identification of isolated half faces in both the upright and inverted orientations was examined as a check of the stimuli (see Appendix A, Table A1). For upright stimuli, mean recognition accuracy for expressions identified in the top half of the face was 92.50% or greater, and in the bottom half of the face was 91.67% or greater. Recognition of inverted stimuli was poorer: top half mean recognition accuracy was 66.67% or greater and in the bottom half of the face, mean accuracy was 89.17% or greater.

Composite effect

Results were analysed separately for expressions in the top and bottom halves as previous studies have found significantly different RTs (Calder et al., 2000; Chiller-Glaus et al., 2011; Palermo et al., 2011) and inconsistent composite effect magnitudes for top and bottom halves (Palermo et al., 2011).

Participants' accuracy scores (percentage correct) were calculated as an average of the four face stimuli for each expression type (see Appendix A, Table A2). RT was considered from the onset of each stimulus for both static and dynamic stimuli. The RT data were calculated for participants' correct responses. With generally high accuracy rates and in line with previous studies, our composite effect analyses focused on RT. Considering the difference in peak expression onset for static and dynamic stimuli, RT data was not directly comparable for static and dynamic stimuli. We therefore calculated baseline corrected scores [(aligned-misaligned)/(aligned+ misaligned)] as a measure of the magnitude of the composite effect for each participant[3] (see Palermo et al., 2011, for a similar procedure with facial

---

[3] See Appendix A, Table A3 for raw RT and baseline RT t-test comparison

expression composite data). Note, a larger baseline corrected RT suggests a stronger composite effect.

To test for composite effects, one-sample *t* tests were conducted comparing mean baseline corrected RT against 0 for each expression. To compare the magnitude of the composite effect for static and dynamic stimuli, separate repeated measures ANOVAs for the top and bottom halves were conducted, including the within-subjects factors of motion (static or dynamic) and expression (top half: anger, sadness, surprise; or bottom half: disgust, fear, joy). Note that any non-integer dfs are the result of Greenhouse-Geisser corrections for violations of sphericity.

Upright Faces

The one sample *t* tests showed significant composite effects for each of the dynamic expression stimuli (all *t*s > 2.72, all *p* < .05). For static expression stimuli, significant composite effects were shown for anger, sadness, disgust, fear, and joy (all *t*s > 2.15, all *p*s < .05). The composite effect for surprise was not significant, $t(29) = .894$, $p = .379$. Data is presented in Figure 2.

Top Half. There was no significant interaction between Motion and Expression, $F(2, 58) = .716$, $p = .493$, $\eta_p^2 = .024$ and no significant main effect of Motion, $F(1, 29) = .032$, $p < .859$, $\eta_p^2 = .001$, indicating similar magnitude composite effects for static and dynamic stimuli. There was a significant main effect of Expression, $F(2, 58) = 4.60$, $p = .014$, $\eta_p^2 = .137$. Overall, composite effects were significantly larger for anger ($M = .05$, $SE = .01$) and sadness ($M = .05$, $SE = .01$) compared to surprise ($M = .02$, $SE = .01$). Anger and sadness did not differ significantly.

Bottom Half. There was no significant interaction between Motion and Expression, $F(2, 58) = 1.52$, $p = .227$, $\eta_p^2 = .050$ and no significant main effect of Expression, $F(2, 58) = 1.58$, $p = .215$, $\eta_p^2 = .052$. There was a significant main effect of Motion, $F(1, 29) = 13.58$, $p = .001$, $\eta_p^2 = .32$, with static stimuli ($M = .05$, $SE = .01$) showing greater composite effects than dynamic stimuli ($M = .02$, $SE = .003$).

INSERT FIGURE 2 HERE

Inverted Faces

For static expression stimuli, there were no significant composite effects for anger, surprise, disgust, fear, and joy (all ts > 1.29, all ps > .05). The composite effect for sadness was significant, $t(29) = 3.47$, $p = .002$, and not different in magnitude to upright sadness ($p = .441$). There were no significant composite effects for the dynamic stimuli surprise, sadness, disgust, and fear (all ts < 1.99, all p > .05). There was a significant composite effect for inverted joy, $t(29) = 3.57$, $p = .001$, that was not different in magnitude to upright joy ($p = .647$). There was a significant composite effect for anger, $t(29) = 2.44$, $p = .021$, however the magnitude of that effect was significantly smaller than for upright anger, $t(29) = 2.38$, $p = .024$.

Discussion

We found clear evidence of holistic processing for dynamic facial expressions (i.e., significant composite effects) regardless of whether the expression was identified in the top or bottom half of a face. These findings extend the results of previous studies using the inversion effect (Ambadar et al., 2005), which suggested holistic coding for subtle dynamic expressions, and are partially in line with Chiller-Glaus et al., (2011), who appeared to find

13

evidence of holistic processing for some dynamic facial expressions. Consistent with Calder et al. (2000) and Palermo et al. (2011), who used a similar paradigm with different static faces, evidence for holistic processing was generally found in both the top and bottom half composites.

In line with the idea that the contribution of holistic and featural information to expression recognition varies with individual emotions (Beaudry, et al, 2014; Bombari et al., 2013; Calvo & Nummennmaa, 2008), comparison of the magnitude of the composite effect across expressions revealed an inconsistent pattern, which depended on whether the expressions were static or dynamic. For expressions that were reliably recognised from the bottom-half of the face (joy, fear and disgust), the composite effect was significantly larger in static than dynamic stimuli. In contrast, for expressions reliably recognised from the top half of the face (anger, sadness and surprise), composite effects for static and dynamic stimuli did not differ except for surprise, where holistic processing was evident for the dynamic but not static stimuli. This finding deviates from Chiller-Glaus, et al (2011), who found composite effects of a similar size for static and dynamic surprise expressions. It is not clear why holistic processing was apparent for dynamic but not static surprise in our study (accuracy and RT were similar to other expressions and isolated halves were well-recognised). However as studies typically collapse results across expressions the generalizability of this finding is unknown. Further research on composite effects in individual expressions would help clarify this issue.

Inverted expressions were tested to check specificity to upright faces and to assess the contribution of the motion signal to the task (since inverted dynamic faces will still display motion signals but are less "face-like"). We found no composite effects when identifying inverted static and dynamic expressions, with the exception of static sadness and dynamic joy and anger. This finding suggests that for most, but not all expressions, holistic processing in

the upright condition is face-specific. The composite effect for inverted dynamic anger was significantly smaller than its upright counterpart, but not for inverted dynamic joy or static sadness. This suggests that for these conditions we cannot attribute performance in their upright counterparts entirely to holistic processing (see McKone et al., 2013). Other factors, such as idiosyncratic motion or emphasis of features, may be facilitating recognition of these expressions. The mouth, for example, is a diagnostic feature in the expression of joy (Bombari et al., 2013; Calvo & Nummennmaa, 2008) and anger is characterized by information in the eye region (Smith, Cottrell, Gosselin & Schyns, 2005). Motion may be helping to disambiguate the diagnostic information in these regions. Overall, while holistic processing for static and dynamic facial expressions is generally specific to upright faces, it appears that the motion signal in expressions of anger and joy and static image factors in expressions of sadness may be contributing other useful cues to expression recognition.

The effect of motion varied when identifying expressions in the top and bottom halves of the face. This could simply be due to the particular expressions that were selected to be presented on the top and bottom. Alternatively, motion from feature changes in the bottom half of a face may be more salient than motion in the top half (Blais, Roy, Fiset, Arguin & Gosselin, 2012). For example, the motion involved in a face changing from neutral to a toothy smile may appear more striking than when eyebrows wrinkle for an angry face. Increased motion salience in particular expressions may enhance, emphasise or disambiguate features and as a result facilitate a greater degree of feature-based processing.

Some studies have shown that motion can improve expression recognition (e.g., Ambadar et al., 2005; Wehrle, Kaiser, Schmidt, & Scherer, 2000) however it is still unclear how it might do so. The results of this study, along with those of Ambadar et al. (2005) and Chiller-Glaus et al. (2012), suggest that the improvement of recognition by motion cannot be attributed to an enhancement of holistic processing, since there was either no difference in the

magnitude of holistic processing between dynamic and static stimuli (top half) or, when identifying expressions in the bottom half of the face, dynamic expressions utilised significantly *less* holistic processing for recognition than static stimuli.

The current study was the first to investigate the role of holistic processing in moving expression composites using only face halves with well-recognised, posed expressions (c.f. Calder et al., 2000; Palermo et al., 2011, with static expressions). The advantage of using clearly identifiable expressions is the presumed reduction of confusion errors. Our results showed that expressions were processed holistically, but that the degree of holistic processing depended on the expression and involvement of motion. However, it is not clear that a similar pattern would emerge for more subtle or less intense expressions. Ambadar et al. (2005) found similar sized inversion effects for static and dynamic expressions, suggesting that the degree of holistic processing was equivalent, at least in subtle static and dynamic expressions. It is possible that the amount of holistic interference from irrelevant halves may depend on the intensity of the expression in each half and this may interact with motion. Chiller-Glaus et al (2010) found composite effects for some expressions with a method that used all combinations of facial expression halves but recognition accuracy was quite low in some cases. And while participants rated emotion intensity in the target half, this data was not analysed across emotion or region. To further explore the conditions under which holistic processing of dynamic facial expressions occurs, future experiments might systematically vary the intensity of the expressions.

Finally, while the composite task is an appropriate measure of holistic processing for dynamic facial expressions, it may not be the best tool for determining relative contributions of holistic and featural processing within individual expressions since: (i) whole face expressions are never presented, (ii) multiple features are contained within a face half, and (iii) different emotions are contained in the top and bottom halves which could be

problematic because some expressions may be processed holistically, some by a single feature and some by multiple features. This latter point is especially important, given that in our study we did not use all possible expression combinations (due primarily to the above mentioned aim of not combining poor- with well-recognised halves, but also because there is evidence that holistic coding magnitude, at least as measured with the identity composite effect, reduces with increasing trials).

Bombari et al., (2013) used the scrambled/blurred method for exploring the contributions of holistic and featural processing of individual static expressions. This method forces participants to adopt either a holistic strategy (using blurred images with little featural information) or a featural processing strategy (using scrambled faces with little holistic information). A similar method may be more appropriate for detailed exploration of holistic and featural processing of individual dynamic expressions, however, dynamic scrambled faces may prove difficult to produce.

To summarise, the use of holistic processing for the recognition of static facial expressions has been demonstrated widely. Here we have provided evidence that holistic processing also occurs for dynamic expressions, validating the large body of existing research on static facial stimuli. Further, our results suggest that the degree of holistic processing depends on the expression to be identified and also on whether expressions are static or dynamic. This is consistent with the idea that the perception of facial expression varies across emotion and is not exclusively holistic or featural but likely to involve some combination of both processes (Beaudry et al., 2014; Tanaka et al., 2012). Overall, the magnitude of holistic processing for dynamic expressions appeared to be either at a similar level to or reduced when compared to static expressions. As motion sometimes improves expression recognition, these findings suggest that motion does not facilitate holistic processing but may instead enhance feature-based processing, at least for some expressions.

References

Ambadar, Z., Schooler, J. W., & Cohn, J. E. (2005). Deciphering the enigmatic face: The importance of facial dynamics in interpreting subtle facial expressions. *Psychological Science, 16,* 403-410. doi: 10.1111/j.0956-7976.2005.01548.x

Bassili, J. N. (1979). Emotion recognition: The role of facial movement and the relative importance of upper and lower areas of the face. *Journal of Personality and Social Psychology, 37,* 2049–2058. doi:10.1037/0022-3514.37.11.2049

Beaudry, O., Roy-Charland, A., Perron, M., Cormier, I., & Tapp, R. (2014). Featural processing in recognition of emotional facial expressions. *Cognition and Emotion, 28,* 416-432. doi:10.1080/02699931.2013.833500

Blais, C., Roy, C., Fiset, D., Arguin, M., & Gosselin, F. (2012). The eyes are not the window to basic emotions. *Neuropsychologia*, *50*(12), 2830-2838.

Bombari, D., Schmid, P. C., Schmid Mast, M., Birri, S., Mast, F. W., & Lobmaier, J. S. (2013). Emotion recognition: the role of featural and configural face information. *The Quarterly Journal of Experimental Psychology, 66,* 2426-2442. doi: 10.1080/17470218.2013.789065

Bruce, V., & Young, A. W. (1986). Understanding face recognition. *British Journal of Psychology, 77,* 305–327. doi: 10.1111/j.2044-8295.1986.tb02199.x

Burton, N., Jeffery, L., Calder, A. J., & Rhodes, G. (2015). How is facial expression coded?. *Journal of Vision*, *15*(1), 1. doi: 10.1167/15.1.1

Calder, A. J., Young, A. W., Keane, J., & Dean, M. (2000). Configural information in facial expression recognition. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 527-551. doi: 10.1037//0096-1523.26.2.527

Calvo, M. G., & Nummenmaa, L. (2008). Detection of emotional faces: salient physical features guide effective visual search. *Journal of Experimental Psychology: General*, *137*(3), 471. doi: 10.1037/a0012771

Chiller-Glaus, S. D., Schwaninger, A., Hofer, F., Kleiner, M., & Knappmeyer, B. (2011). Recognition of emotion in moving and static composite faces. *Swiss Journal of Psychology, 70,* 233-240. doi: 10.1024/1421-0185/a000061

Cohen, J., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behaviour Research Methods, Instruments, and Computers, 25,* 257-271. doi: 10.3758/BF03204507

Cunningham, D., W., & Wallraven, C. (2009). Dynamic information for the recognition of conversational expressions. *Journal of Vision, 9,* 1-17. doi: 10.1167/9.13.7

Ekman, P., & Friesen, W. V. (1978). *Manual for the facial action coding system.* Consulting Psychologists Press.

Ellison, J. W., & Massaro, D. W. (1997). Featural evaluation, integration, and judgement of facial affect. *Journal of Experimental Psychology: Human Perception and Performance, 23,* 213-226

Favelle, S. K., Tobin, A., DeMayo, M., & Palermo, R. (2012). The role of motion in recognising emotional expressions [Abstract]. *Combined Abstracts of 2012 Australian Psychology Conference*, 21

Feldman Barrett, L. (2006). Solving the emotion paradox: categorization and the experience of emotion. *Personality and Social Psychology Review, 10,* 20-46. doi: 10.1207/s15327957pspr1001_2

Fiorentini, P., & Viviani, P. (2011). Is there a dynamic advantage for facial expressions? *Journal of Vision, 11,* 1-15. doi: 10.1167/11.3.17

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in cognitive sciences*, *4*(6), 223-233. doi: 10.1016/S1364-6613(00)01482-0

Kennedy, D. P., & Adolphs, R. (2012). The social brain in psychiatric and neurological disorders. *Trends in cognitive sciences*, *16*, 559-572. doi: doi:10.1016/j.tics.2012.09.006

McKone, E., Davies, A. A., Darke, H., Crookes, K., Wickramariyaratne, T., Zappia, S., …Fernando, D. (2013). Importance of inverted control in measuring holistic face processing with the composite effect and part-whole effect. *Frontiers in Psychology, 4,* 1-21. doi: 10.3389/fpsyg.2013.00033

McKone, E., Martini, P., & Nakayama, K. (2001). Categorical perception of face identity in noise isolates configural processing. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 573-599. doi: 10.1037//0096-1523.27.3.S73

Palermo, R., Willis, M. L., Rivolta, D., McKone, E., Wilson, C. E., & Calder, A. J. (2011). Impaired holistic coding of facial expression and facial identity in congenital prosopagnosia. *Neuropsychologia, 49,* 1226-1235. doi: 10.1016/j.neuropsychologia. 2011.02.021

Piepers, D. W., & Robbins, R. A. (2012). A review and clarification of the terms "holistic," "configural," and "relational" in the face perception literature. *Frontiers in Psychology, 3,* 1-11. doi: 10.3389/fpsyg.2012.00559.r

Pitcher, D., Garrido, L., Walsh, V., & Duchaine, B. C. (2008). Transcranial magnetic stimulation disrupts the perception and embodiment of facial expressions. *The Journal of Neuroscience*, *28*, 8929-8933. doi: 10.1523/JNEUROSCI.1450-08.2008

Rossion, B. (2013). The composite face illusion: a whole window into our understanding of holistic perception. *Visual Cognition, 21,* 139-253. doi: 10.1080/13506285.2013.772929

Skinner, A. L., & Benton, C. P. (2010). Anti-expression aftereffects reveal prototype-referenced coding of facial expressions. *Psychological Science*. *21,* 1248-1253. doi: 10.1177/0956797610380702

Smith, M. L., Cottrell, G. W., Gosselin, F., & Schyns, P. G. (2005). Transmitting and decoding facial expressions. *Psychological Science*, *16*, 184-189. doi: 10.1111/j.0956-7976.2005.00801.x

Tanaka, J. W., Kaiser, M. D., Butler, S., & Le Grand, R. (2012). Mixed emotions: Holistic and analytic perception of facial expressions. *Cognition & Emotion, 26,* 961-977. doi: 10.1080/02699931.2011.630933

van der Schalk, J., Hawk, S. T., Fischer, A. H., & Doosje, B. (2011). Moving faces, looking places: Validation of the Amsterdam dynamic facial expression set (ADFES). *Emotion, 11,* 907-920. doi: 10.1037/a0023853

Wehrle, T., Kaiser, S., Schmidt, S., & Scherer, K. R. (2000). Studying the dynamics of emotional expression using synthesised facial muscle movements. *Journal of Personality and Social Psychology, 78,* 105-119. doi: 10.1037/0022-3514.78.1.105

Appendix A

Table A1
Mean percentage accuracy judgements for isolated half faces (standard deviations)

	Condition			
	Upright		Inverted	
Expression	Static	Dynamic	Static	Dynamic
Bottom				
Disgust	99.17 (4.56)	99.17 (4.56)	99.17 (4.56)	97.50 (10.06)
Fear	91.67 (13.67)	99.17 (4.56)	89.17 (18.20)	97.50 (10.06)
Joy	99.17 (4.56)	100.00 (0.00)	99.17 (4.56)	100.00 (0.00)
Top				
Anger	92.50 (13.37)	93.33 (13.02)	90.00 (14.08)	89.17 (14.21)
Sadness	92.50 (14.58)	97.50 (7.63)	66.67 (27.33)	88.33 (20.48)
Surprise	98.33 (6.34)	98.33 (6.34)	96.67 (10.85)	99.17 (4.56)

Table A2

Mean percentage accuracy of emotion judgements (standard deviations) for static and dynamic stimuli presented in upright and inverted orientations.

Emotion	Static				Dynamic			
	UP-AL	UP-MIS	INV-AL	INV-MIS	UP-AL	UP-MIS	INV-AL	INV-MIS
Bottom								
Disgust	98.89 (3.62)	99.44 (2.11)	97.78 (6.90)	97.78 (5.33)	98.89 (2.88)	99.17 (2.54)	95.00 (9.44)	95.56 (7.81)
Fear	89.17 (11.61)	91.11 (12.93)	93.33 (11.03)	90.83 (16.57)	96.39 (6.45)	95.28 (10.42)	96.67 (6.78)	94.72 (8.61)
Joy	98.61 (3.84)	99.17 (2.54)	98.61 (3.84)	98.06 (4.20)	98.89 (3.62)	98.61 (3.16)	99.72 (1.52)	98.89 (2.88)
Top								
Anger	89.17 (10.53)	90.56 (11.73)	75.56 (12.75)	77.78 (11.44)	89.44 (11.57)	92.22 (10.01)	89.72 (12.12)	87.78 (13.44)
Sadness	96.11 (5.68)	96.39 (7.15)	70.83 (22.50)	78.89 (21.96)	94.72 (9.15)	97.78 (6.54)	92.50 (11.02)	94.72 (8.61)
Surprise	99.72 (1.52)	97.78 (4.86)	98.06 (4.20)	97.50 (5.85)	99.44 (2.11)	98.89 (2.88)	98.89 (2.88)	98.61 (3.16)

Note. UP = Upright, INV = inverted, Al = aligned, MIS = misaligned

Figure Captions

Figure 1. Examples of aligned (left) and misaligned (right) composite stimuli. Recognisable top (anger, surprise, sadness) and bottom (disgust, joy, fear) half expressions were combined to create aligned and misaligned composite stimuli. The task was to recognise the expression on the top (or bottom) half of the face, while ignoring the other half. Holistic coding was measured as quicker recognition of misaligned than aligned composites. Original whole face images taken from the Amsterdam Dynamic Facial Expression Set (Van der Schalk et al., 2011). See supplemental material for examples of dynamic composite expressions.

Figure 2. Baseline corrected RT (for correct responses; [(aligned-misaligned)/(aligned+ misaligned)]) to identify expressions in the top and bottom halves of aligned and misaligned composite static and dynamic stimuli presented in the upright orientation. To account for the difference in peak expression onset for dynamic and static stimuli, baseline corrected RT was calculated from raw RT so the conditions could be directly comparable. Baseline corrected RT ranged from 0.012 - 0.067, with a larger baseline RT indicating a stronger composite effect. Error bars represent +/- one standard error of the mean.