Static and Dynamic Presentation of Emotions in Different Facial Areas: Fear and Surprise Show Influences of Temporal and Spatial Properties*

Holger Hoffmann¹, Harald C. Traue¹, Kerstin Limbrecht-Ecklundt¹, Steffen Walter¹, Henrik Kessler²
¹Medical Psychology, University of Ulm, Ulm, Germany
²Medical Psychology, University of Bonn, Bonn, Germany
Email: holger.hoffmann@uni-ulm.de

Received November 29th, 2012; revised January 6th, 2013; accepted February 6th, 2013

Copyright © 2013 Holger Hoffmann et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

For the presentation of facially expressed emotions in experimental settings a sound knowledge about stimulus properties is pivotal. We hence conducted two experiments to investigate the influence of temporal (static versus dynamic) and spatial (upper versus lower half of the face) properties of facial emotion stimuli on recognition accuracy. In the first experiment, different results were found for the six emotions examined (anger, disgust, fear, happiness, sadness and surprise). Fear and surprise were more accurately recognized when using dynamic stimuli. In the second experiment using only dynamic presentations, recognition rates between upper and lower face varied significantly for most emotions with fear and happiness only being detectable in the upper or lower half respectively. The results suggest an emotion-specific effect for the importance of the facial area.

Keywords: Emotion Recognition; Facial Expressions; Static vs. Dynamic

Introduction

The recognition of emotions is an essential element of social interaction (Buck, 1984; Ekman, 1993). A common paradigm for studying the ability of emotion recognition is the presentation of photographs with different emotional expressions, usually taken at the time of the strongest expression. The disadvantage of this approach is that the ecological validity of such stimuli must be regarded as limited. This is due to the fact that during real-world interactions people must recognize emotions from dynamically changing faces. If one observes emotions over time, it is obvious that they arise at a specific moment, reach their peak and subside again (Onset, Apex, Offset; Hess & Kleck, 2005). In interactions one can also observe that emotions only appear in some areas of the face. The use of static stimuli thus implies the danger of not capturing actual recognition accuracy due to the lack of dynamics and a potential underestimation of the importance of specific facial areas. Hence, it can be concluded that the use of dynamic sequences better reproduces real life situations of emotion recognition and thus allows for more accurate results regarding the recognition of emotions. This could be reflected, for example, in better estimates of recognition accuracy. In the 1980s, Ekman & Friesen (1982) already suspected that static and dynamic stimulus material may result in performance differences.

Another argument in favor of using dynamic stimuli lies in the brain areas being differentially active when presenting static versus dynamic emotional expressions (Trautmann, Fehr, & Herrmann, 2009; Kessler, Doyen-Waldecker, Hofer, Hoffmann, Traue, & Abrler, 2011). Additionally, if natural processes are to be examined, which is not part of this study, dynamic stimuli should be used (e.g., Kilts, Egan, Gideon, Ely, & Hoffman, 2003; LaBar, Crupain, Voyvodic, & McCarthy, 2003; Sato, Kochiyama, Yoshikawa, Naito, & Matsumara, 2004). However, these dynamic stimuli should be standardized, allowing for comparison of results between different studies. The influence of dynamic emotion expressions was already examined in detail in a number of studies with inconsistent results. Harwood, Hall, & Shinkfield (1999) showed, for example, that anger and sadness were more readily recognized when the emotion was presented dynamically. In two classic studies by Bassili (1978, 1979) an improvement in emotion recognition was demonstrated for all emotions when using dynamic stimuli, however. Ambadar, Schooler, & Cohn (2005) used single-static, multi-static, and dynamic stimuli and could demonstrate a robust effect of motion and suggested that this effect was due to the dynamic property of the expression. Additionally, Trautmann, Fehr, & Herrmann (2009) found differences in brain activation patterns comparing the neural processing of static and dynamic stimuli. In their study dynamic stimuli revealed a better recognizability than static stimulus material. A review published in 2013 by Krumhuber, Kappas, & Manstead pronounces the limitation of static stimuli and underline the dynamic nature of...

It is not always the entire face that provides important clues about the expressed emotion. One way of discovering which parts of the face are subjected to a more detailed analysis during the processes of emotion recognition, is by dividing the face into an upper and a lower half and presenting an emotional expression in only one of the areas. Previous studies (Calder, Young, Keane, & Dean, 2000; Bassili, 1978; Bassili, 1979) suggest that emotion recognition is not the same for all basic emotions. This means that there appear to be specific key stimuli for each emotion, which provide the basis for classification. For example, recognition of surprise is associated with observing wide-open eyes. A disguised face, however, is characterized by a wrinkled nose and lifting of the upper lip. It can thus be assumed that the relevance of the specific half of the face depends on the respective basic emotion and its associated key stimuli.

Our study therefore has two aims: 1) Evaluating possible differences between the use of dynamic versus static stimulus material and 2) assessing differential contributions of the upper versus lower half of a facial expression to recognition accuracy.

**General Methods**

The stimuli used for this study were pictures from the JACFEE/JACNeuF (Japanese and Caucasian Facial Expressions of Emotion and Neutral Faces) picture set (Matsumoto & Ekman, 1988). This is a picture set consisting of 56 actors portraying one of seven emotions (anger, contempt, disgust, fear, happiness, sadness and surprise). Half of the actors are male, half female; half are of Japanese and half of Caucasian origin. For our experiments, a subset of 42 actors and six emotions was used, evenly distributed among all sub-sets of stimuli. Contempt was excluded, because this emotion is not considered in the vast majority of studies in the field. In our picture set one actor displays only one emotional and one neutral expression. Several studies have shown the reliability and validity of the JACFEE/JACNeuF picture set in displaying the intended emotions (e.g., Biehl, Matsumoto, Ekman, Hearn, Heider, Kudoh, & Ton, 1997).

The FEMT (Facial Expression Morphing Tool) was used for creating the subtle facial expressions employed in these experiments (Kessler, Hoffmann, Bayerl, Neumann, Basic, Deighton, & Traue, 2005; see also Hoffmann, Kessler, Eppel, Rukavina, & Traue, 2010). This software uses different morphing algorithms to produce intermediate frames between two images. This method was optimized by implementing additional techniques. Sequences were generated using multiple layers that minimized distracting facial information by only morphing the important feature of the face. The use of the multiple layers and special smoothing algorithms allowed us to create realistic transitions from closed to open mouths, for example. The FEMT can generate images in any intensity between 0% (neutral face) and 100% (full-blown emotion). All stimuli used in the study were in color and presented on a computer screen.

**Experiment 1**

In this experiment, the recognition accuracies for static and dynamic stimuli were compared. The hypothesis was that statistically significant differences exist between static and dynamic stimuli and for the six basic emotions.

**Methods**

**Participants**

The study included N = 220 healthy participants. The age of the participants of the experimental group (EG; N = 110) ranged from 18 to 28 years (M = 20.5; SD = 2.0). 70 study participants were female (63.6%), 40 male (36.4%). All subjects of the experimental group gave their written consent to participate in the experiment. A control group (CG; N = 110) was then matched from the FEEL database. The age of the participants in the control group ranged from 19 to 29 years (M = 21.5; SD = 2.3), 63.6% of them were female.

**The FEEL Test (Facially Expressed Emotion Labeling)**

The FEEL test is a computer-based method for measuring individual emotion recognition ability (Kessler, Bayerl, Deighton, & Traue, 2002). It consists of pictures of 42 different actors portraying the six basic emotions (happiness, sadness, disgust, fear, surprise, anger). These images were taken from the JACFEE/JACNeuF picture set (Matsumoto & Ekman, 1988), which was described above. After showing a neutral facial expression, the emotional facial expressions are presented on the computer screen for 300 ms before they must be assigned to a category. For this, a choice box appears from which a selection can be made by clicking on one of the six emotions (forced-choice format). A total of 48 images are presented, as each emotion is shown in a trial run beforehand, so the subjects can acquaint themselves with the task. With a Cronbach’s alpha of r = .77 the test has a high reliability. In the period in which the FEEL test was successfully used, data from 600 healthy subjects of different age, sex and education were collected, so that user-defined control groups can be prepared using this database. Different issues have been examined with this approach (e.g., Hoffmann et al., 2010; Kessler, Roth, von Wietersheim, Deighton, & Traue, 2007).

**Procedure**

The subjects selected from the FEEL database saw static images with a full-blown emotional expression. The experimental group, however, was presented with video sequences that were created from the respective neutral and emotional images, using the FEMT. Since the quality of the picture material only allowed the creation of 36 video sequences, data from the control group were adjusted to account for the missing stimuli. All subjects had to complete six trial runs before the actual test, to ensure familiarity with the procedure. As can be seen in Figure 1, the test procedure was designed to be as identical as possible for the two groups. First, all participants were presented with the neutral expression of an actor. While the control group saw the neutral face for 1500 ms, the experimental group saw it for 1300 ms to 2100 ms. The difference in the presentation time of the neutral face is due to the fact that the dynamic sequences following it had a length of between 400 ms (surprise) and 1200 ms (sadness), depending on the emotion shown. In order to perceive the development of an emotion as natural, a particular temporal sequence must be created. For the emotion of surprise e.g., a much shorter time for the onset is considered as natural compared to the emotion of sadness.
While the experimental group saw the dynamic sequence, the control group was presented with a white screen. Both groups subsequently saw the full-blown emotion for 300 ms, so that for all participants each trial lasted 2500 ms in total. Once the emotional image disappeared from the computer screen, six choice boxes with one emotion label each appeared after 500 ms. Subjects had to choose by mouse click which emotion they had just seen. The participants had ten seconds for deciding before the next trial started. The images with the emotional expressions and the six choice boxes were presented at different times. Presentation of the images was randomized.

### Results

Experiment 1 had a $6 \times 2 \times 2$ mixed design. The within-subject factor was emotion (anger, disgust, fear, happiness, sadness or surprise). The between-subject factors were participants’ gender (female or male) and the type of stimulus material presented (static or dynamic). Analyses were performed using the SPSS 20 software package. Generally, there was no statistical difference in the recognition accuracy between static ($M = 82.5\%$, $SD = 9.3$) and dynamic ($M = 83.7\%$, $SD = 8.2$) stimulus material. Recognition rates for the six emotions differed significantly ($F(5,213) = 47.92; p < .001$) and are shown in Table 1. The interaction between emotion and type of stimulus material was significant, too ($F(5,213) = 5.17; p < .001$), meaning that the type of stimulus material presented influences the recognition accuracy for the six emotions in a different way. We therefore decided to analyze the results looking at the different basic emotions. The results showed that recognition accuracy for surprise ($M_{\text{stat}} = 83.9\%; M_{\text{dyn}} = 90.2\%; p < .01$) and fear ($M_{\text{stat}} = 67.9\%; M_{\text{dyn}} = 75.8\%; p < .05$) increased significantly when presented dynamically as opposed to a static display. In contrast, the recognition accuracy for happiness is statistically higher when using static instead of dynamic stimulus material ($M_{\text{stat}} = 96.4\%; M_{\text{dyn}} = 94.1\%; p < .05$). No significant differences were observed for the emotions anger, disgust and sadness when comparing the experimental and control groups. Female ($M = 83.2\%$) and male ($M = 82.9\%$) participants performed equally.

### Discussion

The conducted experiment partially confirmed the hypothesis that dynamic and static presentations result in significant differences for the individual emotions, although no overall significant difference was found for the use of static and dynamic picture material. The absence of a general effect of display condition over all emotions is consistent with the results of Ambadar et al. (2005), according to which dynamic sequences could not significantly increase recognition accuracy in comparison to a first-last condition.

A differentiated comparison (dynamic versus static) showed that fear and surprise were more readily recognized when the subjects were presented with dynamic sequences. This contradicts the results of a study by Harwood et al. (1999), which reported that the emotions of anger and sadness particularly profited from dynamic presentation. The discrepancy might be explained by the choice of other stimuli. Happiness was less well recognized when presented dynamically, a result that contradicts previous studies. Fear and surprise, on the other hand, tend to be recognized twice as easy when using dynamic sequences. Fear and surprise are often confused, presumably because of the high similarity of the facial expressions (eyes wide open). It appears that these mix-ups can be reduced by means of the additional information provided during the dynamic emergence of the emotion. We assume that the information gained from movement of the mouth and eyes provides particularly important clues for a correct recognition (Jack, Garrod, Yu, Caldara, & Schyns, 2012). The opposite seems to be the case for the recognition of happiness. When quantifying mix-ups, it was striking that in the dynamic presentation condition this emotion was frequently confused with disgust. It is possible that subjects focus on the raising of the lip, which occurs in case of happiness and disgust, so much that other differentiated information, such as wrinkling the nose in case of disgust or the activation of the M. orbicularis oculi in case of happiness, are not considered sufficiently. Contrary to this, other work has shown that spontaneous and deliberate smiles could be distinguished from each other on the basis of dynamic displays, but not static ones (Krumhuber & Manstead, 2009) indicating that the dynamic presentation of happiness increases the ability to distinguish happiness from other emotional states.

Referring to the results of Fiorentini and Viviani (2011), who did not find an advantage for dynamic stimulus material, one should consider the method to develop dynamic stimuli. The authors used high-speed recordings of actors’ facial expressions not morphing sequences of a neutral and a full-blown expression. This may explain the differences in the results and encourage discussing how dynamic stimulus material should be created—with natural expressions or derived from static material. Both options have their (dis-)advantages and cannot be discussed here.

In conclusion, although results are not clear-cut according to Trautmann et al. (2009) dynamic facial expressions might im-

![Figure 1.](Image)

**Figure 1.**
Static vs. dynamic presentation of facial expressions. The left side represents the experimental group (dynamic condition); the right side represents the FEEL group (static condition).
Table 1. Recognition accuracy for the different emotion categories.

<table>
<thead>
<tr>
<th>Emotion</th>
<th>Stimulus type</th>
<th>Static</th>
<th>Dynamic</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anger</td>
<td>91.4 (14.4)</td>
<td>90.5 (14.7)</td>
<td>n.s.</td>
<td></td>
</tr>
<tr>
<td>Disgust</td>
<td>74.4 (25.8)</td>
<td>70.0 (23.7)</td>
<td>n.s.</td>
<td></td>
</tr>
<tr>
<td>Fear</td>
<td>67.9 (25.8)</td>
<td>73.5 (22.9)</td>
<td>p &lt; .05</td>
<td></td>
</tr>
<tr>
<td>Happiness</td>
<td>96.4 (7.3)</td>
<td>94.1 (8.3)</td>
<td>p &lt; .05</td>
<td></td>
</tr>
<tr>
<td>Sadness</td>
<td>80.9 (20.4)</td>
<td>81.8 (20.3)</td>
<td>n.s.</td>
<td></td>
</tr>
<tr>
<td>Surprise</td>
<td>83.9 (15.0)</td>
<td>90.2 (14.0)</td>
<td>p &lt; .01</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>82.5 (9.3)</td>
<td>83.7 (8.2)</td>
<td>n.s.</td>
<td></td>
</tr>
</tbody>
</table>

Note: Standard errors are in parentheses. Recognition accuracy values are in percent.

prove the model how facial expression is processed in humans.

Experiment 2

Emotion recognition accuracy using information from only the lower or the upper part of the face was compared in a second experiment. It was assumed that recognition accuracy differs significantly for these two conditions.

Methods

Participants
The participants were N = 57 students at Ulm University, who gave written consent for participation in the study. Their ages ranged from 18 to 25 years (M = 20.4; SD = 1.6). 42 study participants were female (73.7%), 15 male (26.3%). None of them had been tested in Experiment 1.

Stimuli
In Experiment 2, the facial expressions of the dynamic sequences were synthesized in such a fashion that the transformation was visible only in the upper or the lower half. For this purpose, the face was divided into two halves and a video sequence was generated for each image pair in which the change from a neutral to an emotional expression took place either in the upper or the lower half of the face. This resulted in 72 sequences (6 emotions × 6 sequences × 2 areas of the face). The division of the face into an upper and a lower part was based on the inherent anatomy of the face and can be seen in Figure 2.

Procedure
As in Experiment 1, the subjects first had to complete six trial runs before the actual experiment began. Subsequently they were presented with the 72 sequences in randomized order. Six sequences were presented for each emotion; 50% of the sequences displayed a change in the lower half of the face, and 50% of the sequences displayed a change in the upper half of the face. The course of the trial runs corresponded to that in Experiment 1, and the subsequent evaluation and selection of the emotion shown also followed the same experimental design. One difference between the experiments is the use of a seventh choice box labeled “not recognized” in Experiment 2, which could be selected when the subject was not able to assign the facial expression to one of the six emotions. This was intended to prevent random assignment to an emotion due to lack of choices.

Results
For the statistical analysis we used a generalized linear model considering the dependency structure of our data. Experiment 2 had a 6 × 2 × 2 mixed design with emotion (happiness, disgust, anger, fear, surprise or sadness) and type of stimulus presentation (upper or lower part of the face) as within-subject factors and the gender of the participants (male or female) as a between-subject factor.

Recognition rates differed significantly between the six emotions (Wald $\chi^2 (5, N = 57) = 110.50; p < .001$) and are shown in Table 2. As hypothesized, results show that recognition accuracy differed for the two presentation types, Wald $\chi^2 (1, N = 57) = 29.2; p < .001$. The interaction between emotion and type of stimulus presentation was also significant (Wald $\chi^2 (5, N = 57) = 139.16; p < .001$). Disgust, happiness and sadness were recognized better when the emotional expression was shown in the lower part of the face ($p < .001$). Surprise ($p < .05$) and fear ($p < .001$) in contrast were recognized better when a dynamic change was presented in the upper part of the face. The recognition accuracy of anger seems to be independent of the presentation mode. Subjects were not able to recognize fear and happiness above chance when changes were exclusively presented in the lower half of the face or only in the upper part of the face, respectively. In the case of surprise, on the other hand, high hit rates for both conditions (72.5% in the lower face and 83.0% in the upper face) could be observed. Female and male participants performed equally, no gender effect could be observed.

Discussion
Our hypothesis that presentation of dynamic changes in the upper versus lower part of the face has a differential impact on emotion recognition was partially confirmed. Fear was only reliably detected when the upper half of the face was presented.
In the case of happiness the opposite was true. Surprise was almost equally well recognized in both conditions. With regard to the emotions of happiness and surprise, these results are largely consistent with the studies by Bassili (1978, 1979) and Calder et al. (2000). In the aforementioned studies, surprise was always recognized at a rate of over 70%, regardless of whether the emotional expression was presented in the upper or lower half of the face. Contradictory results are found for anger and sadness. In the present study and the studies by Bassili (1978, 1979) no differences were found for the presentation of anger in the upper or lower part of the face. Calder et al. (2000) showed significantly better recognition accuracy for the presentation of anger in the upper half of the face. In the study by Calder, sadness was recognized when presented in the lower half of the face, in the other studies, however, when presented in the upper half. The differences can possibly be explained by different sample sizes (significantly larger in our study) or the division of the face. While the present research used a facial division based on anatomical features, the abovementioned studies utilized a geometric facial division. Furthermore, in the older studies, only one half of the face was presented, the other half was hidden. This does not conform to real life conditions. The present study showed the entire face. Overall, it could be shown that certain areas of the face are of different relevance for the recognition of basic emotions, though further research is required.

### General Discussion

The results provide new information regarding the question of the ecological validity of stimulus material in the study of emotion recognition. Experiment 1 showed no differences in the use of dynamic and static stimulus material over all emotions, but interesting effects on the level of individual emotions. Experiment 2 showed that the recognition of emotions is differentially influenced by the presentation of the expression in the upper or lower half of the face.

The application of dynamic stimuli is hence not necessary for capturing the assessment of emotion recognition in general. Nevertheless, it appears that dynamic information improves the recognition rates for some emotions. This finding is contradicted by our results regarding the emotion of happiness. Here, higher recognition accuracy was achieved with the use of static stimuli. One must take into account, however, that the recognition rates for this emotion tend to be very high, and that such a result could therefore be due to ceiling effects. Apart from the fact that dynamic stimuli closely represent the natural occurrence of facial emotions, the notion of different brain areas being active when perceiving static versus dynamic stimuli argues in favor of an application of the latter.

Furthermore, emotion recognition appears to depend on the perception of different areas of the face. The information obtained in the relevant half seems to be sufficient for correctly assigning an emotion. This provides opportunities for therapeutic use in people with deficits in the area of emotion recognition.

A study conducted by Adolphs (2002) showed that the recognition accuracy of fear in patients with amygdala lesions could be improved by prompting them to pay attention to the eyes of the presented stimulus. Further research, such as eye-tracking studies, could contribute to the understanding of emotional facial expression analysis.

### Limitations

A limitation of the study concerns the methodological approach. Different times were chosen for the different emotions in the dynamic presentation condition to ensure a natural process. These time specifications, however, are based on the assessment of subjects (Hoffmann et al., 2010). In this experiment the participants were asked to assess dynamic sequences in terms of their realistic representation. It was implicitly assumed that these results reflect the sequence of actual emotion patterns under natural conditions. From an epistemological perspective, this link between perception and production of facial expressions may not necessarily be present.

### REFERENCES


### Table 2.
Recognition accuracy for the different emotion categories.

<table>
<thead>
<tr>
<th>Emotion</th>
<th>Upper</th>
<th>Lower</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anger</td>
<td>59.1 (28.8)</td>
<td>67.3 (35.3)</td>
<td>n.s.</td>
</tr>
<tr>
<td>Disgust</td>
<td>41.5 (32.9)</td>
<td>62.0 (28.5)</td>
<td>(p &lt; .001)</td>
</tr>
<tr>
<td>Fear</td>
<td>50.9 (31.6)</td>
<td>15.2 (28.2)</td>
<td>(p &lt; .001)</td>
</tr>
<tr>
<td>Happiness</td>
<td>18.1 (22.8)</td>
<td>97.1 (11.4)</td>
<td>(p &lt; .001)</td>
</tr>
<tr>
<td>Sadness</td>
<td>43.9 (34.6)</td>
<td>63.7 (27.7)</td>
<td>(p &lt; .001)</td>
</tr>
<tr>
<td>Surprise</td>
<td>83.0 (24.5)</td>
<td>72.5 (29.0)</td>
<td>(p &lt; .05)</td>
</tr>
</tbody>
</table>

Note: Standard errors are in parentheses. Recognition accuracy values are in percent.


