The contribution of different cues of facial movement to the emotional facial expression adaptation aftereffect

Stephan de la Rosa   Max Planck Institute for Biological Cybernetics, Tübingen, Germany   ✉

Martin Giese   Hertie Institute for Clinical Brain Research, Tübingen, Germany   ✉

Heinrich H. Bülthoff   Max Planck Institute for Biological Cybernetics Korea University, Tübingen, Germany   ✉

Cristóbal Curio   Max Planck Institute for Biological Cybernetics, Tübingen, Germany   ✉

Probing emotional facial expression recognition with the adaptation paradigm is one way to investigate the processes underlying emotional face recognition. Previous research suggests that these processes are tuned to dynamic facial information (facial movement). Here we examined the tuning of processes involved in the recognition of emotional facial expressions to different sources of facial movement information. Specifically we investigated the effect of the availability of rigid head movement and intrinsic facial movements (e.g., movement of facial features) on the size of the emotional facial expression adaptation effect. Using a three-dimensional (3D) morphable model that allowed the manipulation of the availability of each of the two factors (intrinsic facial movement, head movement) individually, we examined emotional facial expression adaptation with happy and disgusted faces. Our results show that intrinsic facial movement is necessary for the emergence of an emotional facial expression adaptation effect with dynamic adaptors. The presence of rigid head motion modulates the emotional facial expression adaptation effect only in the presence of intrinsic facial motion. In a second experiment we show these adaptation effects are difficult to explain by merely the perceived intensity and clarity (uniqueness) of the adaptor expressions. Together these results suggest that processes encoding facial expressions are differently tuned to different sources of facial movements.

Introduction

The human face and in particular facial expressions are an important source of information about the cognitive and emotional states of the interaction partner. Importantly facial expressions are inherently dynamic; for example, when a person starts smiling when seeing a friend. In contrast, previous research has mainly investigated the properties of the processes underlying emotional facial expressions by means of static emotional facial expressions (e.g., Benton et al., 2007; Butler, Oruc, Fox, & Barton, 2008; Ellamil, Susskind, & Anderson, 2008; Fox & Barton, 2008; Webster, Kaping, Mizokami, & Duhamel, 2004; Xu, Dayan, Lipkin, & Qian, 2008). Much less is known about the tuning of the processes involved in the

recognition of emotional facial expressions when dynamic information (i.e., facial movement) is present. The present study sought to shed some light onto the tuning of emotional face recognition processes in the presence of dynamic information.

One way to investigate the processes underlying emotional facial expression recognition is by means of visual adaptation. Visual adaptation refers to the perceptual bias that is induced by the prolonged exposure to a stimulus. The resulting adaptation aftereffect allows inferences about the organization of the underlying perceptual processes (Webster, 2011). For example, adaptation to a tilted line causes a subsequently presented vertical line to be perceived tilted to the direction opposite to the adaptor line (Gibson, 1937; Mitchell & Muir, 1976). These results have been explained by means of visual processes that

are tuned to specific orientations and are organized in an interdependent fashion within the visual system. Visual adaptation effects have been demonstrated with simple (e.g., orientation) and complex stimuli (e.g., facial identity and emotional facial expressions) (Butler et al., 2008; Leopold, O'Toole, Vetter, & Blanz, 2001).

Typically the adaptation to one static facial expression (e.g., to a happy face) biases the percept of a subsequently presented static facial expression. In particular the percept of the subsequently presented facial expression is often biased away from the adapted static facial expression (Benton, 2009; Hsu & Young, 2004; Webster, Kaping, Mizokami, & Duhamel 2004). These findings lead to the suggestion that perceptual processes underlying emotional face recognition are organized in a contrastive fashion (Cook, Matei, & Johnston, 2011; Skinner & Benton, 2010). For example, facial expressions might be organized along a dimension that has extreme (peak) expressions at opposite ends of the dimension while a reference (e.g., neutral) facial expression is located in the middle. Adaptation to one facial expression therefore leads to a stronger response of the perceptual processes at the nonadapted side of the dimension compared to the adapted side of the dimension (adaptation aftereffect). When responses are integrated across the entire dimension to form a percept, the resulting perception is biased towards the nonadapted stimulus dimension.

Models of emotional facial expression recognition borrow many of their ideas from other models of face recognition (Giese & Leopold, 2005; Leopold et al., 2001; Valentine, 1991). In particular, models of facial expression assume, like models of facial identity, that facial expressions are represented in a multidimensional framework with a prototype (e.g., Rhodes, Brennan, & Carey, 1987) or a norm-based facial expression (e.g., average facial expression [Valentine, 1991]) at its center. Facial expressions are expressed relative to this prototype or norm-based facial expression (Cook et al., 2011; Skinner & Benton, 2010; Tsao & Freiwald, 2006).

Research on facial expression adaptation focused on the identification of the tuning characteristics of processes involved in the recognition of emotional facial expressions. Several pieces of evidence suggest that perceptual processes involved in emotional face processing are tuned to higher-level visual features that pool visual information across space. Specifically the spatial configuration of low-level facial features (e.g., mouth, eyebrow) is important for inducing an adaptation effect if the facial features themselves are void of emotional expression (Butler et al., 2008; Xu, Dayan, Lipkin, & Qian, 2008). Moreover, emotional facial expression adaption effects have been reported despite large local differences between the adapting and the test stimulus. In particular, adaptation aftereffects have

been reported if the adaptor and test stimulus differed with respect to identity (Ellamil et al., 2008; Fox & Barton, 2007), gender (Fox & Barton, 2007), and viewpoint (Benton et al., 2007). Overall these results suggest that higher order visual features mediate the facial expression adaptation aftereffect and therefore might be important for the recognition of facial expressions (however for local low level adaptation effects in face adaptation see Xu, Dayan, Lipkin, & Qian, 2008).

Much less is known about the process for the integration of specific movement cues in emotional face recognition. There are several reasons as to why encoding of facial movement information (e.g., movement of the eyebrows) might be useful for the recognition of emotional facial expressions. First of all, facial movement information of a facial expression is readily available because facial expressions are inherently dynamic (e.g., when changing from a neutral to a happy facial expression). Secondly, emotional facial expressions are assumed to be universal suggesting that they are carried out (with movement) in similar ways across cultures and therefore movement information is a reliable cue for facial expression recognition (Ekman, 1980; Izard, 1980; Russel, 1994). Hence facial movement information seems to be a readily available and reliable cue for the recognition of a facial expression and its use might be advantageous for the recognition of facial expressions.

To date only one study has addressed the question of whether perceptual processes underlying emotional face recognition are tuned to intrinsic facial movement information. In particular Curio, Giese, Breidt, Kleiner, and Bülthoff (2010) used an adaptation paradigm to show that prolonged visual exposure to dynamic anti-expressions (e.g., antidisgusted face) biases the precept of a subsequently presented ambiguous facial expression. These results suggest that perceptual processes underlying emotional face recognition are tuned to facial motion information.

Several sources of facial movement information exist, for example, head movement (e.g., nodding) and the movement of internal facial features (intrinsic motion) (e.g., eyebrows). Previous research suggested that different sources of facial movement (including head movement) affect the recognition of faces (Bassilli, 1978; Christie & Bruce, 1998; Hill & Johnston, 2001; Knight & Johnston, 1997; Lander & Bruce, 2000) and the recognition of conversational facial expressions (e.g., agreement or surprise) to different amounts (Cunningham & Wallraven, 2009; Nusseck, Cunningham, Wallraven, & Bülthoff, 2008). For example Nusseck et al. (2008) found that freezing the movement of the eye region affected recognition performance of conversational facial expressions less than freezing the movement of the mouth region. Moreover this pattern

depended on the tested conversational facial expression. Hence different parts of the face seem to contribute different amounts to the recognition of conversational facial expressions.

In the present study we sought to examine whether the processes underlying emotional face recognition are tuned to different sources of the facial movement by means of an adaptation paradigm. Specifically we were interested in the effect of rigid head motion and intrinsic facial motion on the emotional face adaptation aftereffect. A modulation of the face adaptation aftereffect by rigid head motion and/or intrinsic facial motion is indicative of the underlying perceptual processes being tuned to this particular kind of facial motion information/cue.

Experiment 1

In an emotional facial expression adaptation paradigm participants repeatedly see an adaptor stimulus that shows one of two facial expressions. In the present experiment this was either a happy or disgusted facial expression (both dynamic and static). Immediately after the presentation of the adaptor participants rated a dynamic test facial expression as either a happy or disgusted face. The dynamic test face's expression varied parametrically between a happy and a disgusted expression.

Because we wanted to selectively manipulate the movement of different parts of the face, we used a novel three-dimensional (3D) computational face model (Curio et al., 2006) for the presentation of facial expressions. This 3D face model showed a face mask that changed its facial expression from a neutral expression to a happy or disgusted facial expression. The facial expressions of the 3D face model were prerecorded facial expressions from a lay actor. The 3D face model allowed us to create new facial expressions (synthesis). The 3D face model gave a parametric control over the movement of individual face features (mouth, head) during synthesis. Moreover it allowed us to morph between two facial expressions. The exact description of the face movement manipulation and the morphing of two facial expressions is described in the Supplementary materials.

The critical manipulation of Experiment 1 consisted of manipulating the availability of rigid head motion and intrinsic facial motion of the adapting stimulus to measure the movement selectivity of the processes underlying emotional face recognition. The factors rigid head movement (on vs. off) and intrinsic facial movement (on vs. off) were completely crossed. The test face was always a dynamic face without rigid head motion. We measured the adaptation aftereffect by

means of psychometric functions relating the happy intensity of the test stimulus to the proportion of happy responses of the participant. Specifically a shift of the psychometric function towards higher happy intensities of the test stimulus in the happy adaptor condition compared to the disgusted adaptor condition would indicate an adaptation effect. If perceptual processes involved in the recognition of emotional facial expressions are tuned to a particular source of face movement, we expect that the availability of this source of information should induce an adaptation aftereffect.

Additionally we also tested the facial expression adaptation aftereffect with static emotional face adaptors. The static adaptors showed a static version of the peak facial expression. In the static adaptor conditions the test stimulus was also a 3D face model with only intrinsic facial movement. The reason for including these control conditions was to compare the size of the adaptation effect between dynamic and static facial expression adaptors. This comparison allows inferences about whether the processes underlying emotional face processing are mainly tuned to static or dynamic information.

Methods

Participants

Ten participants (six male, four females) recruited from the local community of Tübingen participated in the study (mean age: 27.4; $SD = 5.9$). All participants had normal or corrected-to-normal vision. All participants gave their written informed consent prior to the study. The study was conducted in accordance to the Declaration of Helsinki.

Stimuli

The stimulus generation procedure is outlined in detail in the Supplementary materials. In brief the two facial expressions were obtained by recording a happy and disgusted facial expression from a lay actor with a 3D facial motion capture system. We used a 3D morphable face model (Curio et al., 2010) to decompose the recorded facial expressions into rigid head motion and action units activations that described the dynamic changes of the facial form. The action units were chosen similar to the Facial Action Coding System (FACS) proposed by Ekman and Friesen (1978). The action unit activations, as described by vector $\mathbf{w(t)}$, express the activation of action units at a particular point of time relative to a neutral static face. Likewise the head motion transformations, $\mathbf{h(t)}$, describe the head position at a particular point of time relative to head position of a neutral static face. Note that $\mathbf{w(t)}$ provides a mathematical parametrization of internal facial motion over time and $\mathbf{h(t)}$ does the same

for rigid head motion over time. Based on the specified time courses, (w(t), h(t)), a 3D face model can be animated. We calculated **w(t)** and **h(t)** from the prerecorded happy and disgusted facial expressions. In order to morph between happy and disgusted facial expressions, we calculated the weighted sum of the time courses **w(t)** of the happy and the disgusted expression. Because we did not morph between happy and disgusted head motions in the current study, no morphing with respect to **h(t)** was done. Finally in order to turn the internal facial expression on or off action, the unit activation vector **w(t)** was weighted with one or zero, respectively. Likewise a weighting of **h(t)** with one or zero resulted in turning the head movement on or off, respectively. These weighted sums described the morphed facial expression and were used to animate a 3D face model, which was presented to the participants.

Apparatus

All stimuli were presented on a Dell LCD Monitor with a refresh rate of 60 Hz using a Dell desktop computer. The monitor pixel decay time from white to black was between 12 and 15 ms; hence the monitor did not exhibit noticeable image persistence. The presentation software was a custom written software in Matlab using the Psychophysics Toolbox 3 (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007; Pelli, 1997).

Procedure

Participants sat in front of the monitor and received the following information at the beginning of the experiment. After the program started participants were shown a short written summary of the oral instructions on the computer screen ("Please focus on the nose area. Judge the expression after the beep signal: Report whether you saw a happy or disgusted facial expression. Press any key to start"). The participant pressed any key on the keyboard to start the experiment. After the key press the screen turned black for 100 ms and then the first experimental trial started (see Figure 1 for a schematic outline). An experimental trial consisted of four successive adaptor presentations of 1042 ms each, which were separated by an interstimulus interval (showing a black screen) of 100 ms. The fourth adaptor presentation was followed by an interstimulus interval of 200 ms, which was accompanied by a 100 ms 1000 Hz tone that was played over loudspeakers. The tone indicated that the next presented stimulus would be the test stimulus whose facial expression the participant was supposed to judge. The test stimulus was shown for 1042 ms and was replaced by a black screen immediately afterwards. After a 300 ms interstimulus interval the response

screen with the instruction "What did you see? A happy (H) or a disgusted (D) face?" The participants' task was to report whether they perceived the test stimulus as a happy or disgusted expression. A blank screen was presented during the response interval. If participants thought that the test stimulus looked more like a happy expression, they pressed the H key on the keyboard. If they thought that the expression of the test stimulus looked more like a disgust expression, they pressed the D stimulus on the keyboard. The response time was not limited. Moreover no feedback was provided. After the response was given, the next trial started after the participant pressed any key on the keyboard. The experiment consisted of 700 trials. The program stopped every 70th trial displaying a screen that offered participants the chance to take a short break. The participants continued the experiment by pressing any key on the keyboard.

Design

We manipulated the availability of two sources of facial movement of the adapting stimulus, namely rigid head movement (on vs. off) and intrinsic facial movement (on vs. off). Perceptually the four main experimental conditions appeared as follows (see also Supplementary materials for movies of the adaptors). The rigid-head-motion-off/intrinsic-facial motion-off condition is perceptually identical to a static neutral facial expression. The rigid-head-motion-on/intrinsic-facial-motion-off condition appeared as a neutral face doing the head movement of a happy or disgusted expression. The rigid-head-motion-off/intrinsic-facial-motion-on condition looked like a person changing from a neutral to a disgusted or a happy facial expression without moving his/her head. And finally the rigid-head-motion-on/intrinsic-facial-motion-on condition looked like a person normally changing from a neutral to a happy or disgusted expression. Moreover we used two different facial expressions for the adapting stimulus (happy or disgusted). These three factors, rigid head movement, intrinsic facial movement, and facial expression, were completely crossed resulting in eight adaptor conditions.

The test stimulus was a facial expression that was parametrically varied (i.e., morphed) between a happy and a disgusted facial expression. We probed participants' judgments of the test face at seven equidistant points along a linear morph axis (morph weights for the happy expression of 0, 0.217, 0.433, 0.65, 0.87, 1.08, and 1.3). It is important to note that these seven points were equidistant in terms of the morph space but not perceptual space. Although the exact mapping between morph space and perceptual space is unknown (and does not matter for the purposes of the experiment), the facial expression appears as happier with increasing
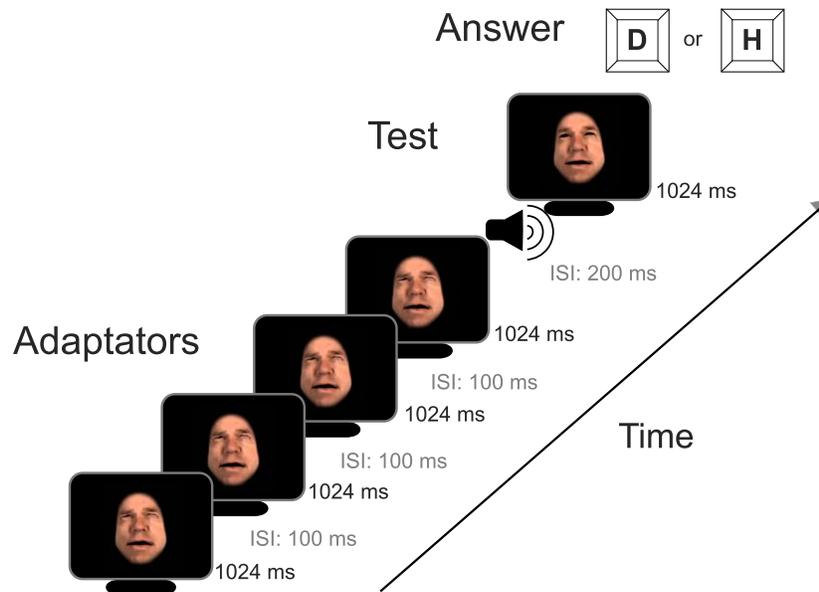
Figure 1. Schematic representation of an experimental trial. The adaptation phase consisted of four adaptor presentations (1042 ms) separated by an ISI (interstimulus interval) of 100 ms. The adaptation and test phase were separated by a 200 ms ISI which was accompanied by a 100 ms 1000 Hz tone. The test stimulus was presented from 1042 ms and participants could give their answer immediately after the presentation by pressing either the H or D key on the keyboard. Shown is a disgusted adaptor and a test stimulus consisting of 70% happy and 30% disgusted.

morph weights. Hence there were seven different test stimuli in total. Each test stimulus was repeated 10 times for each adaptor. Finally the two static adaptors (a happy peak expression adaptor and a disgusted peak expression adaptor) were also combined 10 times with each of the test stimuli. Hence the total number of trials was 7 * 8 * 10 + 7 * 2 * 10 = 700 trials. The trials were presented in random order.

In summary, the main part of the experiment varied four factors: rigid head movement, intrinsic facial movement, facial expression, and morph level. The first three factors manipulated the appearance of the adaptor while the fourth factor manipulated the appearance of the test stimulus. All levels of all factors were tested on every participant resulting in a complete within-subject design. The control condition consisted of a static adaptor and a dynamic test stimulus. The factors facial expression and morph level were also completely crossed for the control conditions and tested in a within-subject manner. The proportion of happy responses served as the dependent variable.

Results

We calculated the percent happy responses for each participant, rigid head motion, intrinsic facial motion, facial expression, and morph level conditions separately for both the main experimental condition as well as

the control conditions. The result for one representative participant is shown in Figure 2.

We then fitted psychometric functions relating the proportion of happy responses to the morph level for each participant, rigid head motion level, intrinsic facial motion level, and facial expression separately. The psychometric functions were fitted in Matlab using a cumulative Weibull function of the form:

$$\text{proportion happy responses} = \gamma + (1 - \gamma - \lambda)\left(1 - e^{-(x/a)^{\beta}}\right), \qquad (1)$$

where gamma is the guessing rate, lambda is the lapse rate, alpha is the location of the psychometric function along the x-axis (i.e., morph level), and beta is the slope of the psychometric function. For the fitting of the psychometric function we fixed gamma at zero and let the three remaining parameters (lambda, alpha, and beta) free to vary. Lambda was allowed to vary between 0% and 5% to account for response lapses, which can significantly affect the quality of the fit (Wichmann & Hill, 2001). The resulting psychometric functions fitted the individual data well (mean $r^2 = 0.994$; $SD_{r\text{-sqr}} = 0.010$).

The resulting average psychometric functions along with the proportion of correct responses are shown in Figure 3. A shift of the psychometric functions along the x-axis between happy and disgusted adaptor conditions is indicative of an adaptation effect. We will first discuss the results of main experimental conditions (black lines) showing dynamic adaptors
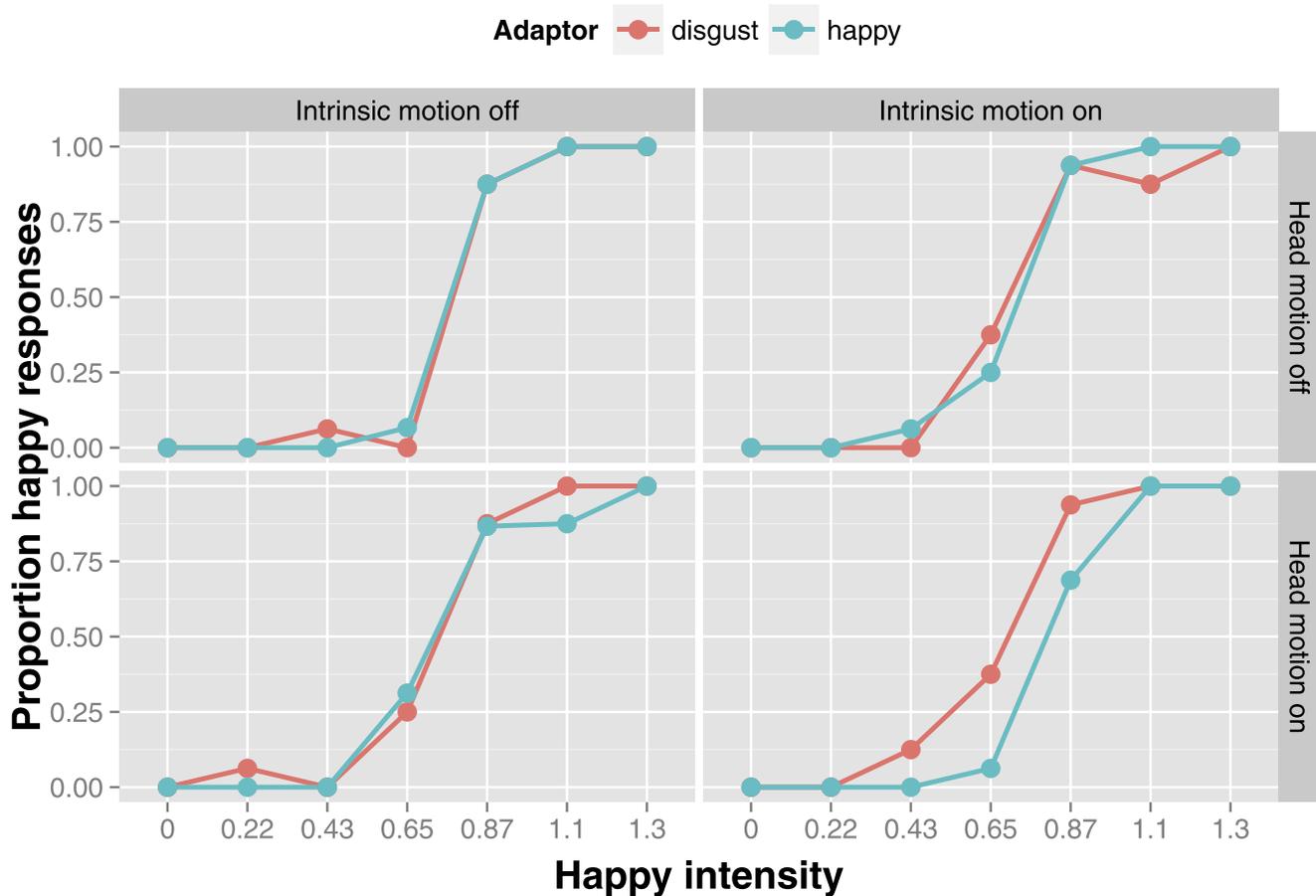
Figure 2. The results of one representative participant of Experiment 1 (only conditions with dynamic adaptors are shown). The panels refer to the different experimental conditions (see top and right label for the experimental condition). Shifts along the x-axis between the blue and red function are suggestive of an adaptation effect.

and later the results of the control conditions (grey lines) showing static adaptors. Note that the adaptor appeared as a static neutral facial expression adaptor in the condition in which both rigid head and intrinsic facial movement were not available (top left panel of Figure 3). In this condition we do not find a shift of the psychometric function suggesting no adaptation occurred with a static neutral adaptor. Adding rigid head movement to a static neutral adaptor (bottom left panel) only induces a minor shift of the psychometric functions in the direction opposite to what previous studies reported. Adding intrinsic facial movement to the static neutral adaptor leads to a somewhat bigger shift of the psychometric functions in the predicted direction (top right panel). That is, the psychometric function for the disgust adaptor is shifted towards lower happy intensities compared to the happy adaptor. Adding both intrinsic facial movement and head movement to a neutral static expression seems to lead to the largest adaptation effect with dynamic adaptors (bottom right panel). Finally the static adaptors showing the peak expression seem to induce

the overall largest adaptation effect (see grey lines in all four panels).

To examine whether the observed differences between psychometric functions in the experimental conditions are statistically significant we looked at the shift of the psychometric functions at the point of subjective equality (PSE). The PSE is the morph level of the test stimulus for which participants are equally likely to give a happy or disgusted response (i.e., 0.5 proportion of happy answers). We determined the PSE for each psychometric function of each participant and experimental condition separately. We submitted the PSE of the main experimental conditions (everything but the control conditions) to a repeated measures ANOVA with rigid head motion, intrinsic facial motion, and facial expression of the adaptor as within-subject factors. Because different PSEs between happy and static adaptor conditions are indicative of an adaptation effect, the factor facial expression of the adaptor measures the adaptation effect. We found a significant main effect of rigid head motion, $F(1, 9) = 5.811$, $p = 0.039$, partial-eta-squared $= 0.395$. The main effects for facial expression of the adaptor, $F(1, 9) =$
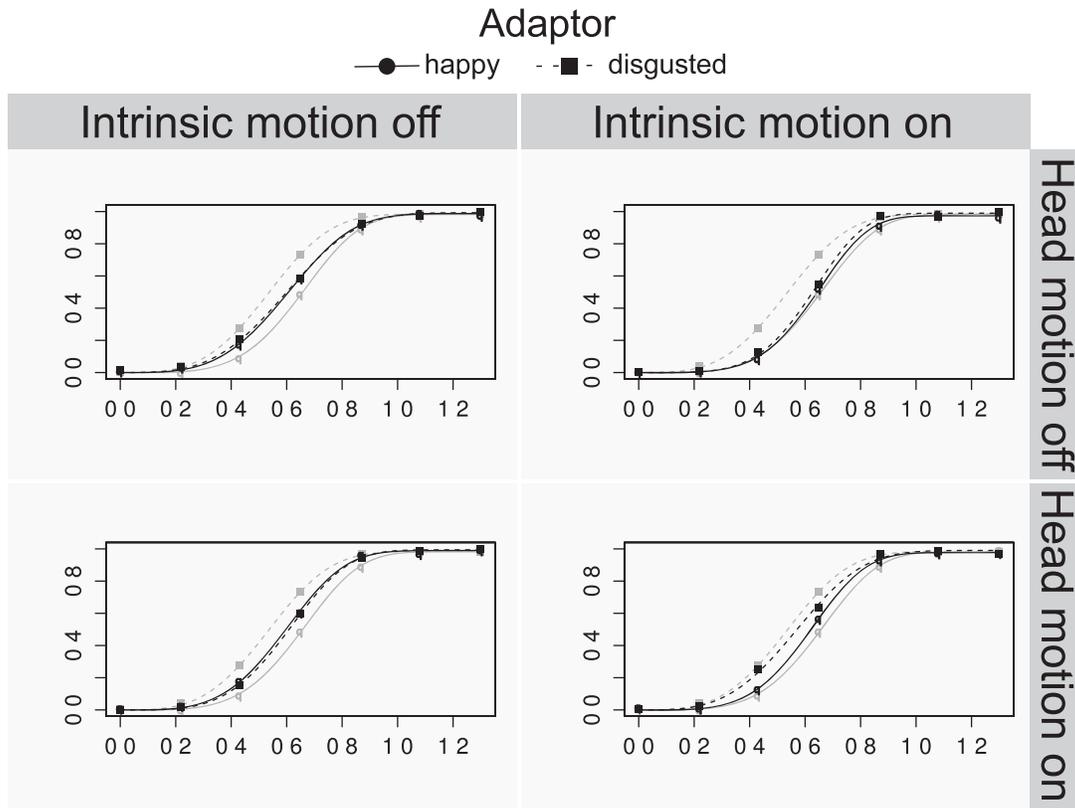
Figure 3. Results of Experiment 1. The panels show the average proportion happy answers as a function of morph level. Each panel refers to a particular combination of rigid head motion and intrinsic facial motion as outlined by the labels to the right and top of the plot. Within each panel the data of the experimental (dynamic adaptor) conditions are plotted in black and the data of the control (static adaptor) conditions are plotted in grey. Note that the control condition was tested only once but for sake of comparison it is plotted within each of the four panels.

3.336, $p = 0.101$, partial-eta-squared $= 0.270$, and intrinsic facial motion, $F(1, 9) = 0.833$, $p = 0.385$, partial-eta-squared $= 0.085$, were not significant. The interaction between facial expression of the adaptor and rigid head motion, $F(1, 9) = 0.011$, $p = 0.918$, partial-eta-squared $= 0.001$, and the interaction between intrinsic facial motion and rigid head motion, $F(1, 9) = 1.922$, $p = 0.199$, partial-eta-squared $= 0.176$, were not significant. The interaction between intrinsic facial motion and facial expression of the adaptor was significant, $F(1, 9) = 5.960$, $p = 0.037$, partial-eta-squared $= 0.398$. Finally the three way interaction between rigid head motion, intrinsic facial motion, and facial expression of the adaptor was significant, $F(1, 9) = 10.165$, $p = 0.011$, partial-eta-squared $= 0.530$. The significant three-way interaction suggests that the effect of happy and disgust adaptors was different for the different combinations of rigid head motion and intrinsic facial motion.

Figure 4 shows the significant three way interaction. We used a Bonferroni corrected paired *t* test to examine for which combinations of rigid head motion and intrinsic facial motion we find a significant shift of the

PSE. All paired *t* tests were evaluated using an adjusted alpha level of 0.01 to accommodate a total of five comparisons (four comparisons for the dynamic adaptors and one for testing one interaction). When both rigid head motion and intrinsic facial motion were not available, the PSE difference between the happy and disgusted adaptor conditions was not significantly different ($M = -0.005$, $SD = 0.039$). We found no significant PSE difference between happy and disgusted adaptors conditions when rigid head motion was available but intrinsic facial motion was unavailable ($M = 0.013$, $SD = 0.039$). We found a significant difference in PSEs when intrinsic facial motion but not rigid head motion was available ($M = -0.030$, $SD = 0.0039$). Finally the PSEs were also significantly different when both rigid head motion and intrinsic facial motion were available ($M = -0.052$, $SD = 0.039$). Hence we only find adaptation effects when the intrinsic facial movement is available.

Does rigid head motion modulate the face adaptation effect when intrinsic facial movement is available (right panel of Figure 4)? We calculated the PSE difference between happy and disgusted adaptor
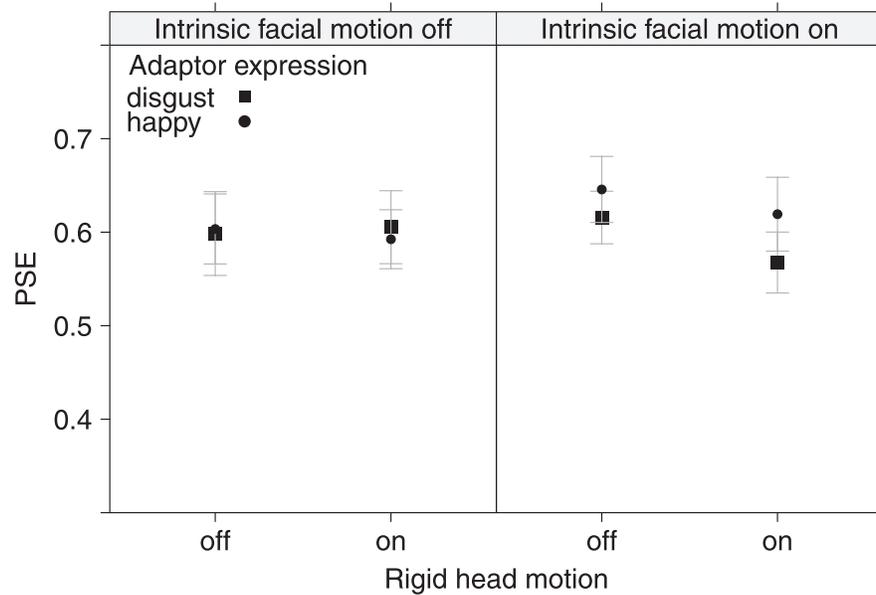
Figure 4. PSE shown for each rigid head motion (along the *x*-axis), intrinsic facial motion (left panel = off; right panel = on), and facial expression (circles = happy adaptors; squares = disgusted adaptors) separately. Bars indicate one standard error from the mean.

conditions for the rigid-head-motion-on and rigid-head-motion-off conditions at the level of intrinsic-facial-motion-on. We then compared the two PSE differences using the above mentioned Bonferroni corrected paired *t* test. We found that the PSE difference was larger in the rigid-head-motion-on condition than in the rigid-head-motion-off condition ($M = 0.021$, $SD = 0.039$). This result suggests that rigid head motion is only able to modulate the facial expression adaptation effect if intrinsic facial movement is available.

The fact that rigid head motion modulates the facial expression adaption effect if intrinsic facial movement is available might seem contradictory to the nonsignificant two-way interaction between head motion and intrinsic facial motion in the main ANOVA above. Remember that the two-way interaction between head and intrinsic facial motion of the main ANOVA, however, does not include the factor facial expression of the adaptor, which measures the adaptation effect. Hence the two-way interaction between head and intrinsic facial movement does not allow inferences about the size of the adaptation effect. It therefore does not contradict the finding that rigid head motion is only able to modulate the facial expression adaptation effect if intrinsic facial movement is available.

Finally we were interested in whether the (facial expression) adaptation aftereffect is stronger for static adaptors than for dynamic adaptors. We compared the adaptation effect associated with adaptors that provide both rigid head motion and intrinsic facial motion with the adaptation effect induced by static peak expression adaptors. An uncorrected paired *t* test revealed that the

adaptation effect was significantly larger with static than with dynamic adaptors, $t(9) = 3.35$, $M_{\text{diff}} = 0.075$, $SD = 0.071$, $p = 0.008$.

In Experiment 1 we found that the size of the adaptation aftereffect depended on the particular combination of available head and intrinsic facial movement cues of the adapting stimulus. The availability of certain facial movement cues is likely to influence the perceived intensity and clarity of a facial expression. As a result, it is possible that in conditions in which the facial expression of the adaptors were perceived as weak or ambiguous due to the lack of certain movement cues, the corresponding facial expression was less adapted compared to a facial expression that was perceived as more intense or clear (i.e., unique). Accordingly one would expect the adaptation effect to be smaller in the first situation. The perceived clarity (uniqueness) and strength of the facial expression of the adapting stimulus might therefore modulate the adaptation aftereffect.

We examined the perceived strength and clarity of the adapting stimulus by means of a questionnaire in Experiment 2. In particular, participants rated the adaptors of Experiment 1 in terms of their emotional content. In addition we explicitly asked participants for clarity and intensity ratings of the displayed facial emotions. If the perceived clarity and strength of the facial expression are behind the observed effects we expect that the perceived clarity and intensity of the adaptors is associated with the magnitude of the adaptation effect. We examined the effect of the perceived clarity and intensity of the emotion on the modulation of the adaptation aftereffect by comparing

participants' ratings of static adaptors to the ratings of dynamic adaptors. If the perceived clarity of the adaptor expression is responsible for the modulation of the adaptation aftereffect, we expect that the expressions of the static adaptors are rated as significantly more intense and clearer than any of the dynamic adaptors.

Experiment 2

Methods

Participants

Fourteen participants of the local community of Tübingen were recruited for the experiment. All participants gave their written informed consent before conducting the study. The study was done in accordance to the Declaration of Helsinki. Participants received 4 Euros for the participation in the study.

Apparatus and stimuli

We used the same adaptors as described in Experiment 1. The questionnaire was programmed in HTML and a browser was used to display the questionnaire. At the top of the page appeared a video showing a facial expression. Below the video the four questions appeared. The questions were as follows. "How happy does the person look like?", "How disgusted does the person look like?", "How intense does the expression of the person look like?", and "How clear is expression of the person?". The order of the four questions was randomly chosen for each adaptor and participant. Below each question a horizontal oriented rating scale consisting of seven radio buttons was displayed. The right and the left end of this rating scale were labeled "not at all" and "completely," respectively. The order in which each of the eight adaptors were probed was randomized across participants.

Procedure

Participants read the written instructions (sent via email) outlining the following procedure. When clicking on the link to the questionnaire in the email, the first page appeared in the browser showing one of the eight adaptors as a video at the top of the page and the four questions with their respective rating scales below it (in random order). The video of the adaptor could be replayed as often as desired by simply clicking on a replay button that appeared at the end of the video. The clarity of the expression should be interpreted as the uniqueness of the expression. Once all eight questions were completed participants had to press the "continue to the next question" button at the end of the page. By clicking this button the next page appeared showing another adaptor along with the four questions. Participants repeated this procedure until they had rated all eight adaptors. The completion of the questionnaire did not take longer than 30 minutes.

Results

We will present the rating results for the question assessing the emotional content and questions assessing the expression intensity separately.

Emotion ratings

The ratings for each question and adaptor movie are shown in Figure 5. For sake of comparison the emotion ratings of the static adaptors are shown as empty circles within each panel. Figure 5 suggests that the emotional ratings of the static adaptors are only clearly higher when the intrinsic facial movement of the adaptor is off. In the other cases static and dynamic adaptors seem to be associated with similar ratings. A Bonferroni corrected t test (corrected for the four comparisons within each panel in Figure 5) showed that only the dynamic disgusted adaptor was rated as significantly less disgusted than the static disgusted adaptor in the conditions where the intrinsic facial movement was off. No other statistical significant differences were found.

Intensity and clarity ratings

The intensity and clarity ratings of the adaptor movies are shown in Figure 6 for the head and intrinsic facial movement manipulation separately. The static adaptor ratings are indicated by the empty circles. A comparison of the adaptor ratings of the dynamic conditions with the static adaptor ratings showed that participants rated the static adaptors as more intense and clearer than dynamic adaptors when the dynamic adaptor was lacking intrinsic facial movement. A Bonferroni corrected t test (corrected for the four comparisons within each panel in Figure 6) showed that the static disgusted adaptor was perceived as significantly more intense and clear than the dynamic adaptor that was void of intrinsic facial movement and head movement. On the other hand, the dynamic happy adaptor that contained both head and intrinsic facial movement was rated as significantly clearer than the static adaptor.

In summary participants perceive the emotions conveyed by static and dynamic adaptors as similar when the expressions contain intrinsic facial movement.
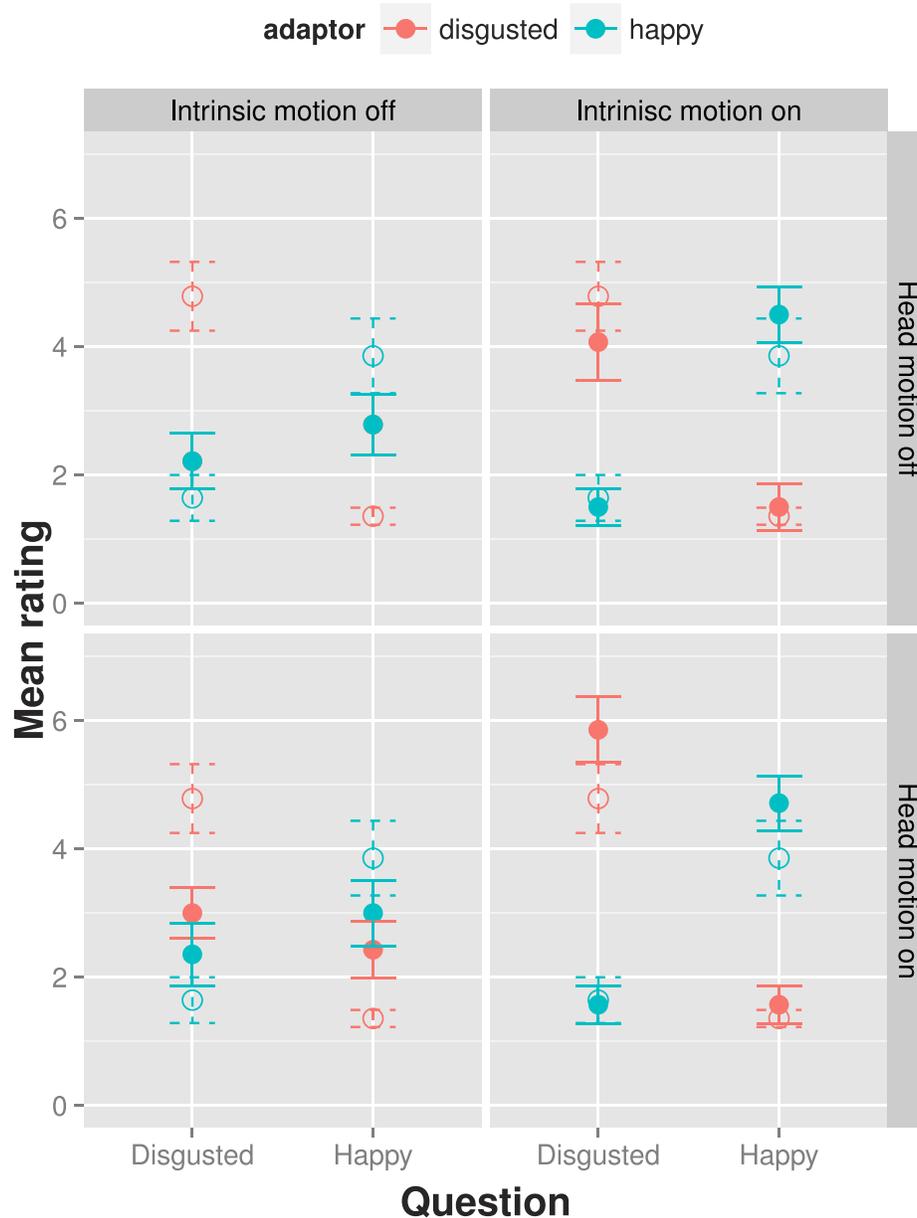
Figure 5. Emotion rating results of Experiment 2. The mean emotion ratings are shown for the adaptors for each head and intrinsic facial movement condition (across panels) and facial expression (different colors) separately. Bars indicate one standard error (*SE*) from the mean. The empty circles and dashed lines refer to the mean ratings and *SE*, respectively, of the static adaptors; they are shown, for sake of comparison, within each panel.

Even more, when both head and intrinsic facial movement are available, participants perceive the happy dynamic adaptor as significantly clearer than its static counterpart. Overall there seems to be little evidence for participants perceiving the emotions conveyed by dynamic adaptors generally as less intense and clear than static adaptors.

These results are difficult to reconcile with the idea that the adaptation aftereffects are modulated by perceived intensity and clarity of the adaptor expression. If this were the case, we would expect the clarity and intensity ratings of the static adaptors to be the largest since the largest adaptation effects were found with these adaptors. Our results indicate, however, that dynamic adaptors with intrinsic facial movement were similarly rated to static adaptors and in one instance even received a significantly higher clarity rating than their static counterparts. We therefore suggest that the perceived intensity of the adaptor's facial expression is not the only factor behind the modulation of the observed adaptation effects in Experiment 1.
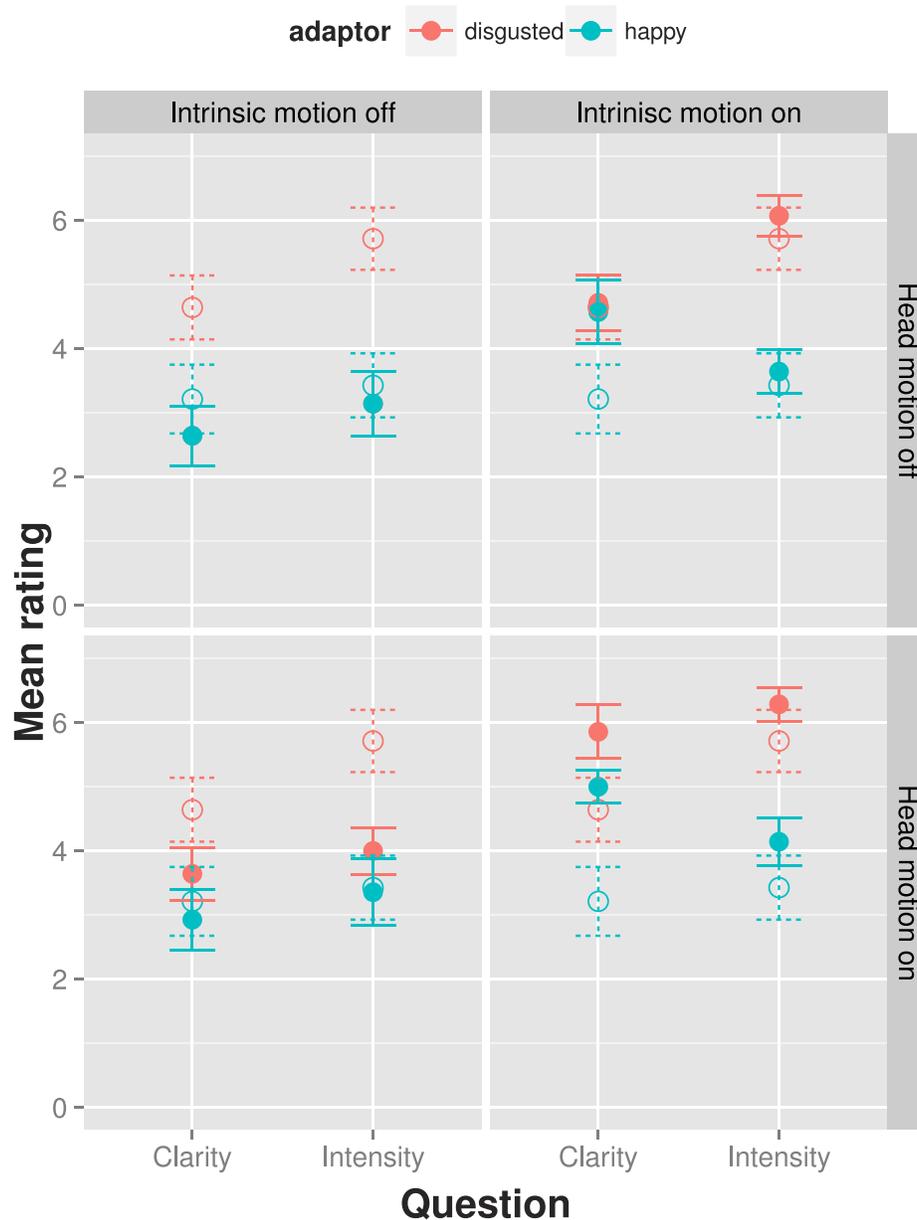
Figure 6. Clarity and intensity rating results of Experiment 2. The mean clarity and intensity ratings are shown for the adaptors for each head and intrinsic facial movement condition (across panels) and facial expression (different colors) separately. Bars indicate one *SE* from the mean. The empty circles and dashed lines refer to the mean ratings and *SE*, respectively, of the static adaptors. They are shown, for sake of comparison, within each panel.

Discussion

The aim of the present study was to examine the contribution of different types of facial motion to the facial expression adaptation aftereffect. We used an adaptation paradigm together with a novel 3D face model to address this question. The application of the 3D face animation model (Curio et al., 2006) allowed a parametric manipulation of different sources of facial movement in terms of time and the parametric morphing of the test stimulus between happy and disgusted facial expressions on the level of facial action units, inspired by FACS (Facial Action Coding System; Ekman & Friesen, 1978). In the experiment we varied the availability of the rigid head motion information and intrinsic facial motion information of the adaptor using a completely crossed within-subject design and measured the adaptation aftereffect. Facial expressions in this study changed from a neutral expression to either a happy or disgusted expression. Additionally we also probed the adaptation aftereffect with static peak emotional expressions (static happy or static disgusted face). Our results show that the adaptation effect is

strongest when static peak emotional expressions are used as adaptors. The second strongest adaptation effect is found with dynamic facial expression adaptors containing both rigid head motion and intrinsic facial motion. A significantly smaller adaptation effect is found for dynamic adaptors that provide only the intrinsic facial motion but no rigid head motion. Adaptors providing rigid head motion alone or a neutral facial expression did not induce an adaptation effect.

In a second experiment we examined whether the perceived intensity and clarity (uniqueness) of the adaptor expression modulates the adaptation effects in Experiment 1. If this were the case we expect that the perceived intensity and clarity of the facial expression would be largest in the static condition since the adaptation was largest in this condition in Experiment 1. In contrast to this assumption we found that intensity and clarity ratings were similar for dynamic adaptors with intrinsic facial movement and static adaptors. In one case the clarity of the dynamic adaptor expression was rated significantly clearer than the corresponding static expression. Hence we suggest that the perceived intensity and clarity of the facial expressions are not the major contributing factor to the modulation of the adaptation effects in Experiment 1.

Can our results be explained by an adaptation effect that is merely driven by static emotional facial expressions? For example, note that only the last movie frame(s) of the dynamic facial expression show the peak facial expression. Hence one could argue that the peak facial expression drives the adaptation aftereffect and not the intrinsic facial movement. The smaller adaptation effect with dynamic adaptors could then be explained in terms of a shorter adaptation period when dynamic adaptors are used. Specifically one could argue that the adaptation is reduced due to the availability of the peak facial expression only during the last movie frames with dynamic adaptors. We think that this explanation is unlikely to account for the observed effects for several reasons. First, if adaptation effects are solely driven by the peak facial expressions, one would expect that the adaptation effect is the same for all experimental conditions in which the peak facial expression is available (i.e., conditions in which intrinsic facial motion is available). However our results show that the size of the adaptation effect in the intrinsic facial motion conditions depends on the availability of rigid head motion (whose presentation is not confounded with the presentation of a peak expression). Hence facial motion information is able to modulate the size of the adaptation effect in the absence of a static peak facial expression. Moreover, according to the hypothesis that only peak facial expressions are responsible for the adaptation effect, adaptation aftereffects for dynamic adaptors should be always smaller than for static adaptors when facial

expression are changing from neutral to the probed facial expression. Curio et al. (2010) however found emotional facial expression aftereffects with dynamic facial expressions to be of similar magnitude than those with static adaptors using emotional anti-expression as adaptors. Importantly these adaptors also changed from a neutral expression to the anti-expression. Taken together we think that it is unlikely that the adaptation effects are merely driven by the peak facial expression component of a dynamic facial expression.

Is it possible that rigid head motion induces less of an adaptation effect than intrinsic facial movement because the test stimulus contains intrinsic facial movement but no rigid head movement? In other words, one could argue that the diminished effect of adaptation with rigid head motion is simply owed to the absence of this cue in the test stimulus. According to this explanation only the intrinsic facial movement adaptor should induce the largest adaptation effect. This is however not what we found. The largest adaptation was found with static adaptors that show the peak frame but are void of both intrinsic facial movement and rigid head movement. This result demonstrates that adaptation is induced even if the adaptor stimulus does not share the same cues with the test stimulus. Hence we do not think that diminished effect of rigid is due to a lack of rigid head motion in the test stimulus.

Wu, Xu, Dayan, and Qian (2009) reported face adaptation aftereffects with facial motion that were induced by the similarity of the backgrounds in which the faces are presented. We do not think that a similar effect might explain the adaptation differences between static and dynamic adaptors or between different dynamic adaptors because our background was void of any texture in all conditions.

Can the results be explained in terms of the recognition model proposed by Giese and Poggio (2003)? In this model the output of snapshot neurons—each tuned to a particular static body posture—feeds into motion pattern neurons which mediates the recognition of a biological motion. According to this model the size of the adaptation effect increases with an increase of the number of adapted snapshot neurons feeding into the same motion pattern neuron. In the experiment reported here, the dynamic adaptor shares many more physical similarities with the dynamic test expression than the static adaptors. Therefore dynamic adaptor expressions should activate more snapshot neurons than static adaptor expressions and as a result adaptation to dynamic expressions should induce larger adaptation effects than adaptation to static expressions. Because our results show the opposite, the findings are therefore difficult to explain in terms of the Giese and Poggio (2003) model.

It is possible that the advantage of static adaptors with respect to inducing a face expression aftereffect is

expression specific. The ability to recognize basic emotions across cultures even when they are presented statically points to a special property of these facial expressions, namely, that the static information is sufficient for their recognition. Other facial expressions, however, seem to rely to a much larger degree on dynamic visual information (Nusseck et al. 2008). Hence the importance of dynamic visual information in the recognition of a facial expression might be expression specific. An interesting future research avenue would be to examine the effect of dynamic information on the facial expression aftereffect using other nonbasic emotions whose recognition seems to rely more on dynamic visual information.

The finding that both static and dynamic adaptors induced an adaptation effect can be explained by emotional face recognition processes being tuned to both static and dynamic facial information. The idea that visual recognition might be tuned to both static and dynamic visual information is not novel and has been proposed for the motion aftereffect (Hiris & Blake, 1992; Mather, Pavan, Campana, & Casco, 2008; Verstraten, Frederiksen, Van Wezel, Lankheet, & Van de Grind, 1996). More research is necessary to disentangle the interplay of dynamic and static visual information for facial expression recognition. The finding that static adaptors induced the largest adaptation effects suggests that emotional facial recognition processes seem to rely mainly on static facial information.

In summary we showed that different sources of dynamic face information affected the adaptation aftereffect to different degrees. Specifically the movement of intrinsic facial features seems to be critical for the presence of an adaptation aftereffect. However the largest adaptation aftereffects were found with static adaptors. We therefore suggest that processes underlying emotional facial expression recognition are mainly tuned to static face information although they are also tuned to dynamic face information.

Acknowledgments

Commercial relationships: none.

Corresponding authors: Stephan de la Rosa, Heinrich Bülthoff, or Cristobal Curio.
Email: delarosa@kyb.tuebingen.mpg.de; heinrich.buelthoff@tuebingen.mpg.de; cristobal.curio@tuebingen.mpg.de.
Address: Department of Human Perception, Cognition and Action, Max Planck Institute for Biological Cybernetics, Tübingen, Germany.

References

Bassilli, J. N. (1978). Facial motion in the perception of faces and emotional expressions. *Journal of Experimental Psychology: Human Perception and Performance, 4*(3), 373–379.

Benton, C. P. (2009). Effect of photographic negation on face expression aftereffects. *Perception, 38*(9), 1267–1274, doi:10.1068/p6468.

Benton, C. P., Etchells, P. J., Porter, G., Clark, A. P., Penton-Voak, I. S., & Nikolov, S. G. (2007). Turning the other cheek: The viewpoint dependence of facial expression after-effects. *Proceedings of the Royal Society B: Biological Sciences, 274*(1622), 2131–2137, doi:10.1098/rspb.2007.0473.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*(4), 433–436.

Butler, A., Oruc, I., Fox, C. J., & Barton, J. J. S. (2008). Factors contributing to the adaptation aftereffects of facial expression. *Brain Research, 1191,* 116–126, doi:10.1016/j.brainres.2007.10.101.

Christie, F., & Bruce, V. (1998). The role of dynamic information in the recognition of unfamiliar faces. *Memory & Cognition, 26*(4), 780–790.

Cook, R., Matei, M., & Johnston, A. (2011). Exploring expression space: Adaptation to orthogonal and anti-expressions. *Journal of Vision, 11*(4):2, 1–9, http://www.journalofvision.org/content/11/4/2, doi:10.1167/11.4.2. [PubMed] [Article]

Cunningham, D. W., & Wallraven, C. (2009). Dynamic information for the recognition of conversational expressions. *Journal of Vision, 9*(13):7, 1–17, http://www.journalofvision.org/content/9/13/7, doi:10.1167/9.13.7. [PubMed] [Article]

Curio, C., Breidt, M., Kleiner, M., Vuong, Q. C., Giese, M. A., & Bülthoff, H. H. (2006, July). Semantic 3D motion retargeting for facial animation. In R. W. Fleming & S. Kim (Eds.): Proceedings of the 3rd Symposium on Applied Perception in Graphics and Visualization (pp. 77–84), Boston, MA.

Curio, C., Giese, M. A., Breidt, M., Kleiner, M., &

Bülthoff, H. H. (2010). Recognition of dynamic facial action probed by visual sdaptation. In C. Curio, H. H. Bülthoff, & M. A. Giese (Eds.), *Dynamic faces: Insights from experiments and computation* (pp. 47–65). Cambridge, MA: MIT Press.

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

Ekman, P., & Friesen, W. (1978). *Facial action coding system: A technique for the measurement of facial movement.* Palo Alto, CA: Consulting Psychologists Press.

Ellamil, M., Susskind, J. M., & Anderson, A. K. (2008). Examinations of identity invariance in facial expression adaptation. *Cognitive, Affective, & Behavioral Neuroscience, 8*(3), 273–281, doi:10.3758/CABN.8.3.273.

Fox, C. J., & Barton, J. J. S. (2007). What is adapted in face adaptation? The neural representations of expression in the human visual system. *Brain Research, 1127*(1), 80–89, doi:10.1016/j.brainres.2006.09.104.

Fox, C. J., & Barton, J. J. S. (2008). It doesn't matter how you feel. The facial identity aftereffect is invariant to changes in facial expression. *Journal of Vision, 8*(3):11, 1–13. http://www.journalofvision.org/content/8/3/11, doi:10.1167/8.3.11. [PubMed] [Article]

Gibson, J. J. (1937). Adaptation, after-effect, and contrast in the perception of tilted lines. II. Simultaneous contrast and the areal restriction of the after-effect. *Journal of Experimental Psychology, 20*(6), 553–569.

Giese, M. A., & Leopold, D. A. (2005). Physiologically inspired neural model for the encoding of face spaces. *Neurocomputing, 65–66,* 93–101, doi:10.1016/j.neucom.2004.10.060.

Giese, M. A., & Poggio, T. (2003). Neural mechanisms for the recognition of biological movements. *Nature Reviews Neuroscience, 4*(3), 179–192, doi:10.1038/nrn1057.

Hill, H., & Johnston, A. (2001). Categorizing sex and identity from the biological motion of faces. *Current Biology: CB, 11*(11), 880–885.

Hiris, E., & Blake, R. (1992). Another perspective on the visual motion aftereffect. *Proceedings of the National Academy of Sciences of the United States of America, 89*(19), 9025–9028.

Hsu, S.-M., & Young, A. W. (2004). Adaptation effects in facial expression recognition. *Visual Cognition, 11,* 871–899.

Izard, C. E. (1980). Cross-cultural perspectives on emotion and emotion communication. In W. Triandis (Ed.), *Handbook of cross-cultural psychology* (pp. 185–220). Boston: Allyn & Bacon.

Kleiner, M., Brainard, D., & Pelli, D. G. (2007). What's new in Psychtoolbox-3? *Perception, 36,* 14.

Knight, B., & Johnston, A. (1997). The role of movement in face recognition. *Visual Cognition, 4*(3), 265–273, doi:10.1080/713756764.

Lander, K., & Bruce, V. (2000). Recognizing famous faces: Exploring the benefits of facial motion. *Ecological Psychology, 12*(4), 259–272, doi:10.1207/S15326969ECO1204_01.

Leopold, D. A., O'Toole, A. J., Vetter, T., & Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level aftereffects. *Nature Neuroscience, 4*(1), 89–94, doi:10.1038/82947.

Mather, G., Pavan, A., Campana, G., & Casco, C. (2008). The motion aftereffect reloaded. *Trends in Cognitive Sciences, 12*(12), 481–487, doi:10.1016/j.tics.2008.09.002.

Mitchell, D. E., & Muir, D. W. (1976). Does the tilt after-effect occur in the oblique meridian? *Vision Research, 16*(6), 609–613, doi:10.1016/0042-6989(76)90007-9.

Nusseck, M., Cunningham, D. W., Wallraven, C., & Bülthoff, H. H. (2008). The contribution of different facial regions to the recognition of conversational expressions. *Journal of Vision, 8*(8):1, 1–23, http://www.journalofvision.org/content/8/8/1, doi: 10.1167/8.8.1. [PubMed] [Article]

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*(4), 437–442.

Rhodes, G., Brennan, S., & Carey, S. (1987). Identification and ratings of caricatures: Implications for mental representations of faces. *Cognitive Psychology, 19*(4), 473–497.

Russel, J. A. (1994). Is there universal recognition of emotion from facial expression. *Psychological Bulletin, 115*(1), 102–141.

Skinner, A. L., & Benton, C. P. (2010). Anti-expression aftereffects reveal prototype-referenced coding of facial expressions. *Psychological Science, 21*(9), 1248–1253, doi:10.1177/0956797610380702.

Tsao, D. Y., & Freiwald, W. A. (2006). What's so special about the average face? *Trends in Cognitive Sciences, 10*(9), 391–393, doi:10.1016/j.tics.2006.07.009.

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *The Quarterly Journal of Experimental Psychology Section A, 43*(2), 161–204, doi:10.1080/14640749108400966.

Verstraten, F. A., Fredericksen, R. E., Van Wezel, R. J., Lankheet, M. J., & Van de Grind, W. A. (1996). Recovery from adaptation for dynamic and static motion aftereffects: Evidence for two mechanisms. *Vision Research, 36*(3), 421–424.

Webster, M. A. (2011). Adaptation and visual coding. *Journal of Vision, 11*(5):3, 1–23, http://www.journalofvision.org/content/11/5/3, doi:10.1167/11.5.3. [PubMed] [Article]

Webster, M. A., Kaping, D., Mizokami, Y., & Duhamel, P. (2004). Adaptation to natural facial categories. *Nature, 428*(April), 357–360, doi:10.1038/nature02361.1.

Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception & Psychophysics, 63*(8), 1293–1313.

Wu, J., Xu, H., Dayan, P., & Qian, N. (2009). The role of background statistics in face adaptation. *The Journal of Neuroscience, 39*, 12035–12044.

Xu, H., Dayan, P., Lipkin, R. M., & Qian, N. (2008). Adaptation across the cortical hierarchy: Low-level curve adaptation affects high-level facial-expression judgments. *The Journal of Neuroscience, 28*(13), 3374–3383, doi:10.1523/JNEUROSCI.0182-08.2008.