

The efficiency of dynamic and static facial expression recognition

Jason M. Gold

Department of Psychological and Brain Sciences,
Indiana University, Bloomington, Indiana, USA



Jarrett D. Barker

Department of Psychological and Brain Sciences,
Indiana University, Bloomington, Indiana, USA



Shawn Barr

Department of Psychological and Brain Sciences,
Indiana University, Bloomington, Indiana, USA



Jennifer L. Bittner

Department of Psychological and Brain Sciences,
Indiana University, Bloomington, Indiana, USA



W. Drew Bromfield

Department of Psychological and Brain Sciences,
Indiana University, Bloomington, Indiana, USA



Nicole Chu

Department of Psychological and Brain Sciences,
Indiana University, Bloomington, Indiana, USA



Roy A. Goode

Department of Psychological and Brain Sciences,
Indiana University, Bloomington, Indiana, USA



Doori Lee

Department of Psychological and Brain Sciences,
Indiana University, Bloomington, Indiana, USA



Michael Simmons

Department of Psychological and Brain Sciences,
Indiana University, Bloomington, Indiana, USA



Aparna Srinath

Department of Psychological and Brain Sciences,
Indiana University, Bloomington, Indiana, USA



Unlike frozen snapshots of facial expressions that we often see in photographs, natural facial expressions are dynamic events that unfold in a particular fashion over time. But how important are the temporal properties of expressions for our ability to reliably extract information about a person's emotional state? We addressed this question experimentally by gauging human performance in recognizing facial expressions with varying temporal properties relative to that of a statistically optimal ("ideal") observer. We found that people recognized emotions just as efficiently when viewing them as naturally evolving dynamic events, temporally reversed events, temporally randomized events, or single images

frozen in time. Our results suggest that the dynamic properties of human facial movements may play a surprisingly small role in people's ability to infer the emotional states of others from their facial expressions.

Introduction

The ability to reliably identify another person's emotional state from their body language is an integral part of even the most basic of human social interac-

Citation: Gold, J. M., Barker, J. D., Barr, S., Bittner, J. L., Bromfield, W. D., Chu, N., et al. (2013). The efficiency of dynamic and static facial expression recognition. *Journal of Vision*, 13(5):23, 1–12, <http://www.journalofvision.org/content/13/5/23>, doi:10.1167/13.5.23.

tions. Previous research has shown that we rely upon a wide variety of information sources in order to infer others' emotions, such as their facial expressions (Ekman, 1993), gait (Clarke, Bradshaw, Field, Hampson, & Rose, 2005), and vocal prosody (Banse & Scherer, 1996). Although we are clearly able to derive information about emotion through static images of facial expressions, it is less clear whether we gain any perceptual advantage from viewing facial expressions as naturally evolving dynamic events. The results of previous experiments that have attempted to address this question have been inconsistent, with some suggesting dynamic cues may offer significant processing advantages (Ambadar, Schooler, & Cohn, 2005; Bassili, 1979; Bould & Morris, 2008; Brainard, 1997; Cunningham & Wallraven, 2009; Edwards, 1998; Knappmeyer, Thornton, & Bulthoff, 2003; Lander, Christie, & Bruce, 1999; Wehrle, Kaiser, Schmidt, & Scherer, 2000) and others suggesting they may not (Bruce et al., 1999; Christie & Bruce, 1998; Fiorentini & Viviani, 2011; Kamachi et al., 2001; Katsyri & Sams, 2008). Thus, it currently remains unclear what role dynamic cues play in our ability to extract information about emotion from facial expressions (see Fiorentini & Viviani, 2011, for a recent overview).

One factor that may be at least partially responsible for this lack of consistency across studies is the absence of an objective measure of how much information is available to observers across varying tasks and stimulus conditions. For example, several studies have directly compared observers' ability to recognize expressions that dynamically evolve over time to their performance with only single static images of fully evolved expressions (Ambadar et al., 2005; Bassili 1979; Bould & Morris, 2008; Chiller-Glaus, Schwaninger, Hofer, Kleiner, & Knappmeyer, 2011; Christie & Bruce, 1998; Cunningham & Wallraven, 2009; Fiorentini & Viviani, 2011; Fujimura & Suzuki, 2010; Katsyri & Sams, 2008; Lander & Bruce, 2000; O'Toole et al., 2011). In most cases, this kind of direct comparison of human performance across different stimulus conditions can be challenging because it confounds the ability to use information with the physical availability of information. That is, differences in performance across two or more stimulus conditions (e.g., dynamic vs. static expressions) can be due to either differences in the ability of observers to make use of available information, differences in the physical availability of information, or some combination of the two. As a result, it is difficult to interpret such data in the absence of an objective measure of stimulus information.

One approach to this problem that has been used successfully in other contexts is to measure the performance of an "ideal observer"—a statistically optimal decision rule that is guaranteed to yield an upper bound on task performance (Geisler, 1989, 2004;

Green & Swets, 1966). Because the ideal observer is only limited by the physical availability of information, its performance provides a direct measure of the relative amount of information available across a set of conditions. Further, comparison of human to ideal performance, a measure known as "efficiency," factors out the effects of any variations in information content, yielding a pure measure of a human observer's relative ability to make use of available information across a set of conditions (Geisler, 2011).

Given the above, the goals of the experiment reported here were to use ideal observer analysis to (a) measure the amount of information physically available to an observer when recognizing static and dynamic facial expressions, and (b) measure how efficiently human observers make use of that information when performing those tasks. Specifically, we measured both human and ideal observer contrast energy thresholds for recognizing the emotions of facial expressions that were presented as either naturally unfolding dynamic events, temporally randomized dynamic events, or single static snapshots frozen in time. Using this approach, we were able to objectively compare the amount of information carried by dynamic, static, and temporally randomized expressions, as well as measure the relative amount of information used by human observers when viewing each kind of event.

Methods

Observers

Sixteen observers (eight male, eight female) participated in the experiment. Half of the participants were authors and the other half were naive to the purposes of the experiment. Naive subjects participated in the experiment for course credit. All observers had normal or corrected-to-normal visual acuity.

Stimuli and apparatus

We generated our stimuli by digitally recording (at 30 frames/s) a set of eight actors (four male, four female) who were asked to make six different basic facial expressions of emotion (Ekman, 1972). The emotions were happiness, sadness, anger, disgust, fear, and surprise. Each expression started from a neutral state and naturally evolved into a full expression ("Dynamic" expressions; Figure 1). The transition point between a neutral expression and the initial formation of an expression, as well as the point at which an expression reached its fully articulated state, was determined by two raters (both authors). Most expressions were fully

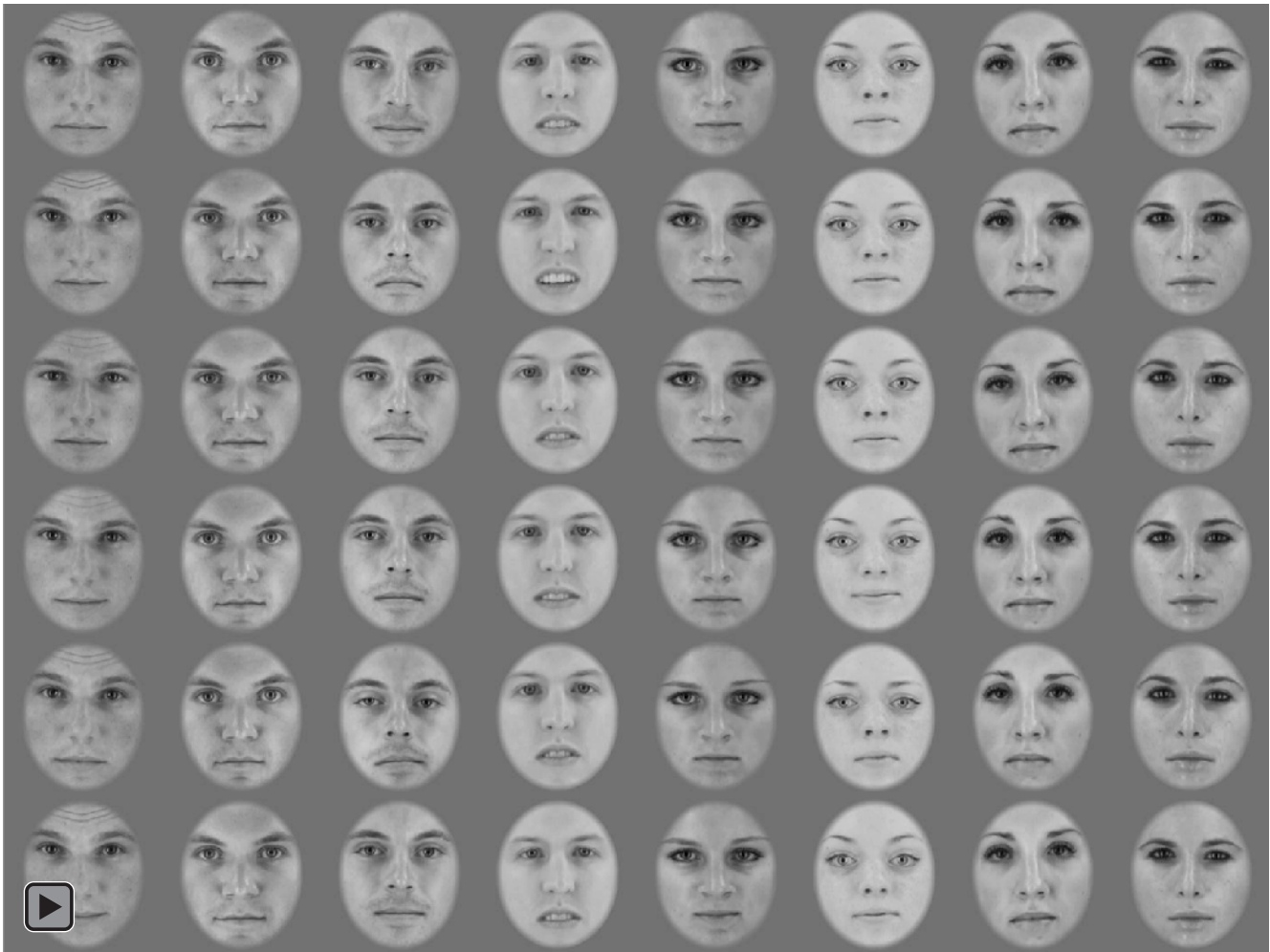


Figure 1. The full Dynamic movies of each actor (columns) making each facial expression (rows, from top to bottom: anger, disgust, fear, happiness, sadness, surprise).

formed over the course of 30 frames, and none required more than 30 frames to reach their fruition. Actors were asked to maintain their full expressions for several seconds, and thus most of the movies that required less than 30 frames to reach their apex included additional fully expressed frames. Due to head movements or other artifacts, the apex frame of a small number of movies was repeated for the remainder of the 30 frames (this occurred for only 7 out of 48 movies, and involved repeating the last frame between 1–4 times).

For each movie frame, the actor's face was resized and centered within an oval aperture of a fixed height/width ratio (164×140 pixels, $2.7^\circ \times 2.2^\circ$ from a viewing distance of 130 cm). Each movie was expressed in units of contrast by scaling it between -1 and 1 . The oval aperture was then centered within a 256×256 pixel ($4.2^\circ \times 4.2^\circ$) uniform gray background of zero contrast (which corresponded to a luminance of 38 cd/m^2). The aperture edge was slightly blurred to produce a gradual transition between the face image and the gray

background. The frame rate of the CRT on which the stimuli were displayed was 85 Hz, and each of the 30 individual movie images was repeated for three successive screen refreshes, producing a stimulus duration of 1059 ms at an effective frame rate of 30 frames/s. The contrast energy (i.e., integrated squared contrast) of the stimuli was adjusted across trials in order to obtain contrast energy expression discrimination thresholds in each condition (see "Threshold and efficiency measurement" below for more details on the staircase procedure and threshold measurement). The contrast energy of the stimulus was adjusted by multiplying a selected movie by an appropriate constant and then converting the pixel contrast values to luminance values according to the equation $L_{\text{pix}} = C_{\text{pix}} * L_{\text{bg}} + L_{\text{bg}}$, where L_{pix} is pixel luminance, C_{pix} is pixel contrast and L_{bg} is background luminance. These luminance values were then converted to RGB values (according to a look-up table generated using a Minolta LS-100 photometer [Konica Minolta Sensing, Inc., Ramsey, NJ]) and displayed on the CRT.

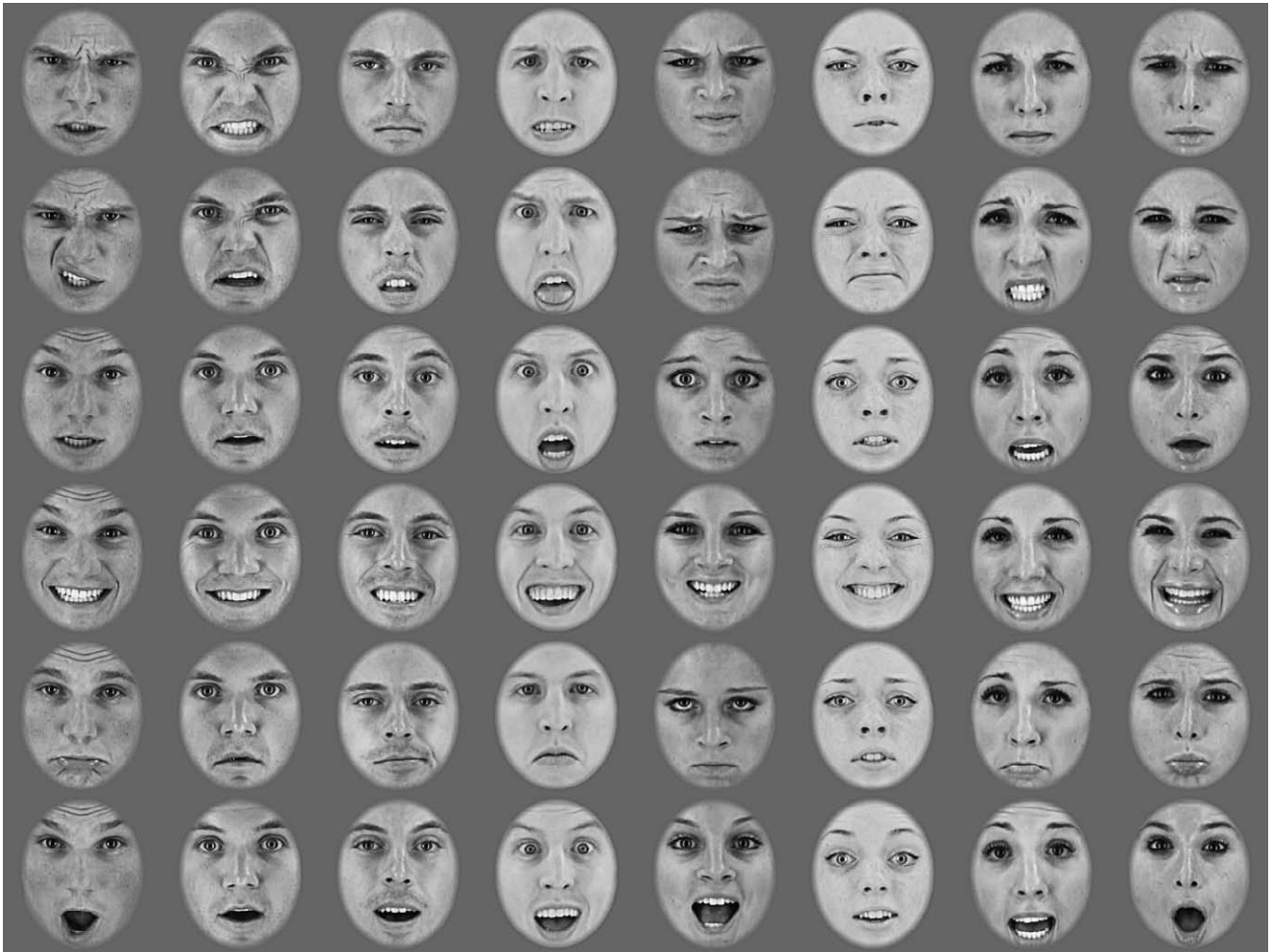


Figure 2. The Static images (i.e., final movie frames) of each actor (columns) making each facial expression (rows, from top to bottom: anger, disgust, fear, happiness, sadness, surprise).

From these Dynamic expression movies, we generated two additional sets of stimuli: a set of “Static” expressions (Figure 2), in which the final image of each Dynamic expression was shown for the same duration as the full Dynamic expression movie; and a set of “Shuffled” expressions (Figure 3), in which the order of the images that made up each Dynamic movie was randomly permuted.

In the Static condition, movies were generated by replicating the last frame from each corresponding Dynamic movie 30 times (Figure 2). In the Shuffled condition, a set of 10 different random frame permutations was generated and these orders were applied to each of the Dynamic movies, producing a total of 480 temporally shuffled movies (see the “Ideal observer” section below for more on the use of a limited set of random permutations).

To make the task more difficult and to carry out the ideal observer analysis, we added a unique sample of Gaussian white contrast noise ($\sigma = 0.25$) to each pixel of

the stimulus on each trial. The noise matched the movie both in size and mean luminance, and was added to the stimuli in all three conditions (Dynamic, Static, and Shuffled). Previous experiments have shown that contrast energy thresholds in similar tasks increase linearly with external noise variance (e.g., Gold, Bennett, & Sekuler, 1999b; Tjan, Braje, Legge, & Kersten, 1995), a property that implies observers are using a “contrast invariant” strategy (i.e., their strategy is independent of the contrast of the signal and the variance of the externally added noise; Pelli, 1990). Thus, it is unlikely that the manipulation of contrast and the addition of external noise in our task had a significant influence on observers’ recognition strategies.

Procedure

On each trial, a movie was randomly chosen with equal probability and displayed on the computer

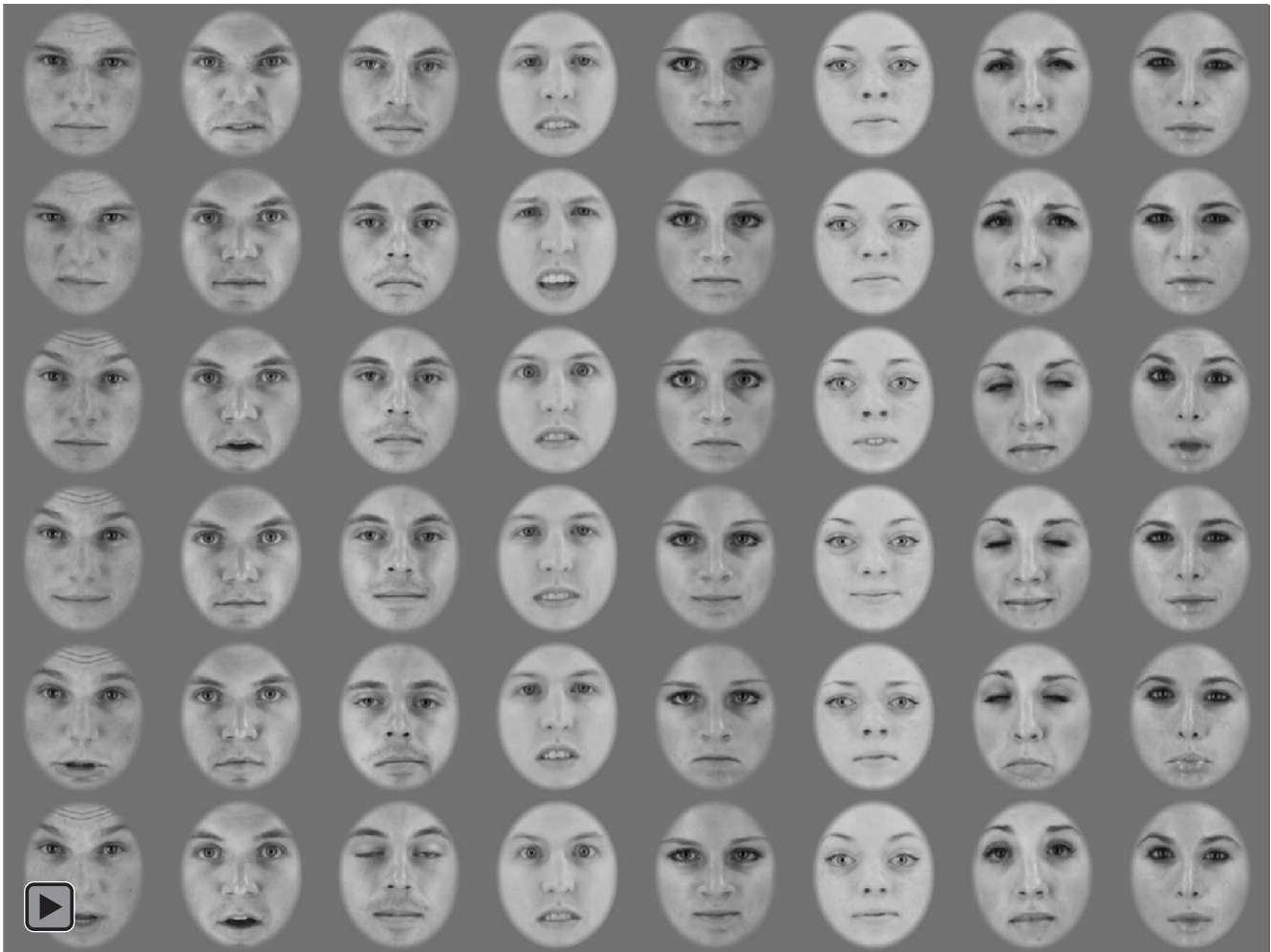


Figure 3. An example of a random frame permutation (‘Shuffle’) applied to each dynamic movie shown in Figure 1.

screen. After each stimulus presentation, a selection window with the names of the six different possible emotions appeared and remained on the screen for an unlimited amount of time until the participant chose one of the options with a mouse click. Auditory accuracy feedback was given after each trial. Each human observer participated in all three conditions of the experiment. Trials were blocked by condition, and the order of conditions was randomized across participants. Each observer was familiarized with the stimuli from all three experimental conditions prior to participating in the experiment by completing 15 practice trials in each condition.

Threshold and efficiency measurement

Task difficulty was manipulated by varying the contrast energy (i.e., integrated squared contrast) of the movies across 240 trials in each condition using a 1-down, 1-up adaptive staircase procedure (Macmillan &

Creelman, 1991). Weibull psychometric functions were fit to the staircase data, and threshold was defined as the contrast energy that yielded 50% correct performance (note that chance performance would be $\sim 17\%$ correct in a 1 of 6 classification task; 50% correct was defined as threshold because it falls very close to the middle of a typical psychometric function in a 1 of 6 recognition task). Efficiency was defined as the ratio of ideal-to-human contrast energy threshold in a given condition.

Ideal observer

The ideal observer’s performance was measured by computer simulations in each stimulus condition that the human observers were tested in (Dynamic, Static, and Shuffled). The ideal decision rule for our task and stimuli was derived using Bayes’ rule and is similar in principle to other tasks involving 1-of-N recognition (e.g., Gold, Bennett, & Sekuler, 1999a; Tjan et al.,

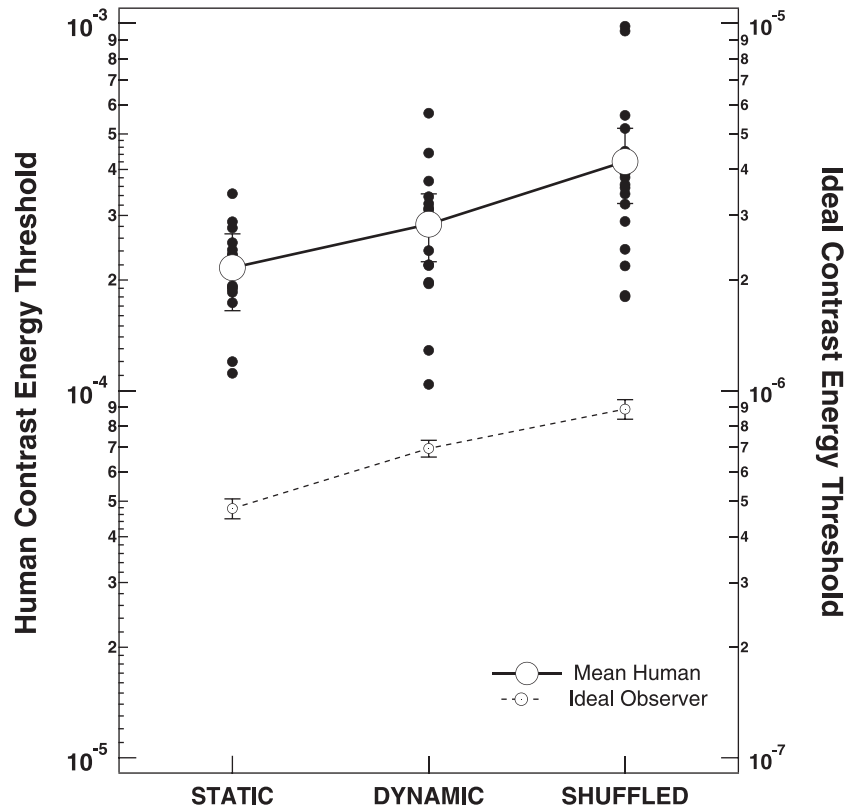


Figure 4. Contrast energy thresholds for human observers (small closed circles, large open circles and solid line, left ordinate) and the ideal observer (small open circles and dashed line, right ordinate) for the Static, Dynamic, and Shuffled facial expression stimuli. Each solid circle corresponds to an individual human participant; large open circles correspond to the mean threshold across participants in each condition. Error bars on each mean correspond to ± 1 SEM.

1995). In our experiment, observers were asked to determine the expression, E_i (where i refers to the i^{th} of r possible expressions), that was most likely to have appeared within the noisy stimulus data, D . According to Bayes' rule, the a posteriori probability of E_i having been presented given D can be expressed as

$$P(E_i|D) = P(E_i)P(D|E_i)/P(D) \quad (1)$$

For our task and stimuli, the prior probability of seeing any given expression, $P(E_i)$, and the normalizing factor $P(D)$ are both constant for all E_i , and thus can be removed without changing the relative ordering of $P(E_i|D)$. Therefore, the ideal observer chooses the expression that maximizes $P(D|E_i)$. For the case where there are m possible faces for each expression shown in additive Gaussian white noise, the ideal observer must compute this probability for all m possible faces within each expression category (all of which are equally probable) and compute the summed probability across faces, resulting in the following probability function:

$$P(D|E_i) = \sum_{k=1}^m \prod_{j=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(D_j - E_{ijk})^2} \quad (2)$$

where n is the total number of pixels in the entire

stimulus (i.e., all pixels of all 30 frames) and σ is the standard deviation of the Gaussian distribution from which the external noise was generated. The ideal decision rule is to choose the expression E_i that maximizes this function.

In the case of the Static and Dynamic conditions, $m = 8$ faces within each expression category. In the case of the Shuffled condition, it is intractable to carry out the ideal observer analysis on all possible randomly permuted frame orders. Therefore, we generated a subset of p random frame permutation orders, where $p = 10$, and applied this random order to each face, yielding 8 (faces) \times 10 (frame orders) = 80 faces within each expression category. This produced an extra layer of summation in the probability function described in Equation 2, corresponding to the number of permutation orders, i.e.:

$$P(D|E_i) = \sum_{l=1}^p \sum_{k=1}^m \prod_{j=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(D_j - E_{ijk})^2} \quad (3)$$

Thresholds for the ideal observer in all conditions were measured in the same fashion as the human observers, with 2,500 simulated trials per condition.

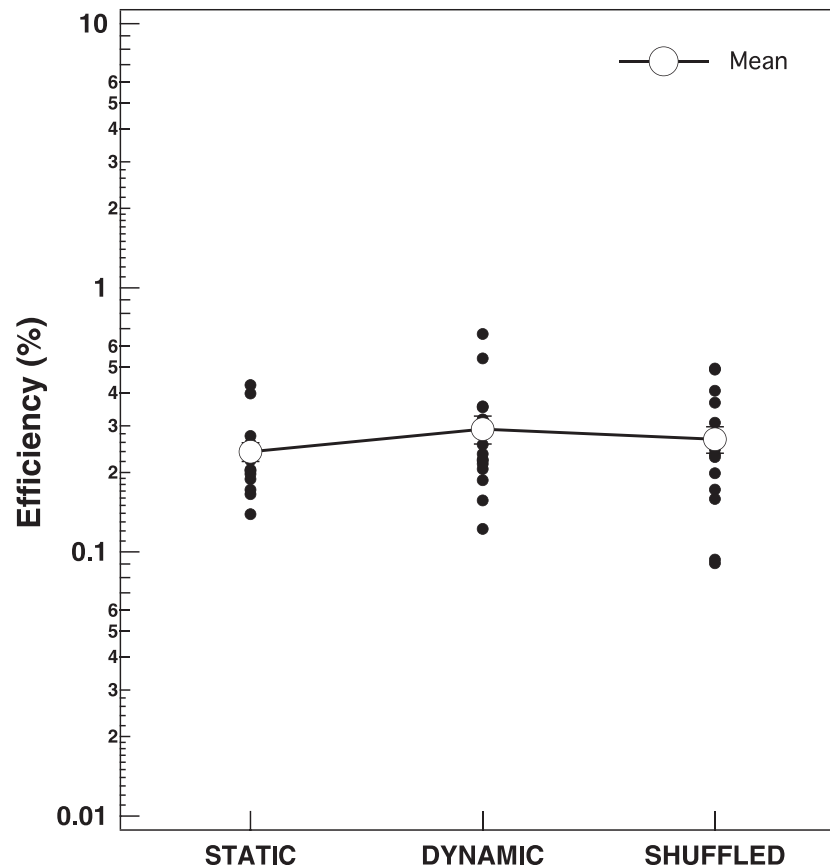


Figure 5. Corresponding efficiencies (ideal/human contrast energy thresholds) for each human participant shown in Figure 4. Small closed circles correspond to efficiencies for individual observers; large open circles correspond to the mean efficiency across observers. Error bars on each mean correspond to ± 1 SEM.

Results

The results of our experiment are shown in Figures 4 and 5. Figure 4 plots thresholds for each individual observer (small solid circles) and the mean across observers (large open circles) in each condition. These data show that average performance was best for Static expressions, slightly worse for Dynamic expressions, and worst for Shuffled expressions. A one-way repeated measures analysis of variance showed there was a significant effect of stimulus type, $F(2, 30) = 10.73$, $p = 0.0003$, and a post-hoc Tukey honestly significant difference test indicated this effect was due to thresholds being significantly greater for the Shuffled than both the Static ($p = 0.001$) and the Dynamic ($p = 0.046$) conditions; the difference between Dynamic and Static thresholds was not statistically significant ($p = 0.23$).

The dashed line and small open circles in Figure 4 plot the thresholds for the ideal observer in each condition of our experiment. Note that human thresholds are plotted on the left axis and ideal

thresholds are plotted on the right axis. Both axes span a total of two log units, but the ideal axis is shifted two log units lower than the human axis. Separate axes were used in this fashion because the ideal observer was over two orders of magnitude better than our human observers in all conditions, as is typical with most complex pattern recognition tasks (e.g., Gold et al., 1999a; Gold, Tadin, Cook, & Blake, 2008; Tjan et al., 1995). However, because we were interested in observers' relative performance across conditions rather than their absolute performance, different axes were used to place the human and ideal data in a similar range.

The results of the ideal observer analysis show that the amount of discriminative information was not the same for Dynamic, Static, and Shuffled expressions. In fact, the ideal observer's pattern of thresholds was remarkably similar to what we obtained with our human participants: The ideal observer performed best with Static faces, worse with Dynamic faces, and worst with Shuffled faces. The resulting human efficiencies, computed as the ratio of ideal/human threshold in each condition for each human observer, are plotted in Figure 5. These data show that, when the amount of physically

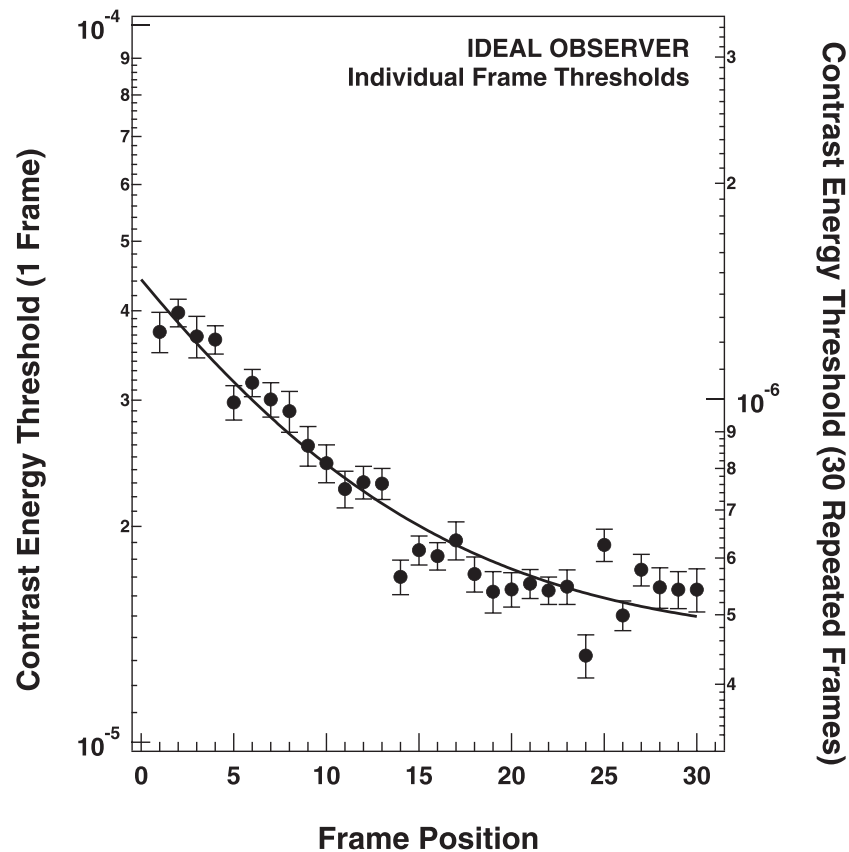


Figure 6. Ideal observer contrast energy thresholds for individual frames in the Dynamic facial expressions. The left-hand scale plots thresholds for each frame shown a single time in Gaussian White noise. The right-hand scale plots thresholds for each frame repeated 30 times in Gaussian White noise. The solid line is the best fitting exponential function of the form $y = y_0 + \alpha e^{-(x-x_0)/\tau}$. Error bars for each threshold correspond to ± 1 SD, and were estimated by bootstrap simulations (Efron & Tibshirani, 1993).

available information is taken into account, efficiency is nearly identical for Dynamic, Static, and Shuffled expressions. A one-way repeated measures analysis of variance confirmed that there were no significant differences across conditions, $F(2, 30) = 1.39$, $p = 0.265$.

Discussion

The results of our experiment offer some interesting new insights into the information carried by static and dynamic facial expressions as well as the human ability to make use of that information. First and foremost, our finding that efficiency did not differ significantly for Static and Dynamic expressions indicates that when the amount of physically available information is taken into account, the presence of dynamic facial cues appears to offer no additional information processing benefit to human observers beyond those carried by a single static snapshot of a fully expressed emotion.

Second, our ideal observer analysis revealed that a static snapshot of a fully articulated facial expression actually carries more discriminative information with respect to discerning others' emotional states than a naturally evolving dynamic event. At first, this may seem like a counterintuitive result. However, recall that our Static face stimuli were generated by replicating the final frame of our Dynamic face stimuli. If we were to assume that these final frames tend to be the most distinctive exemplars from each Dynamic expression, then it would follow that our Static stimuli should in turn carry more information overall, as each frame of the Static stimuli would be maximally informative across time. Note that repeating the same movie frame 30 times in a row does add additional information beyond that carried by a single frame shown only once. This is because each repeated frame is presented in a new sample of noise, and thus the ideal observer (and potentially *any* observer) benefits from receiving repeated presentations of the same signal due to the statistical averaging of noise.

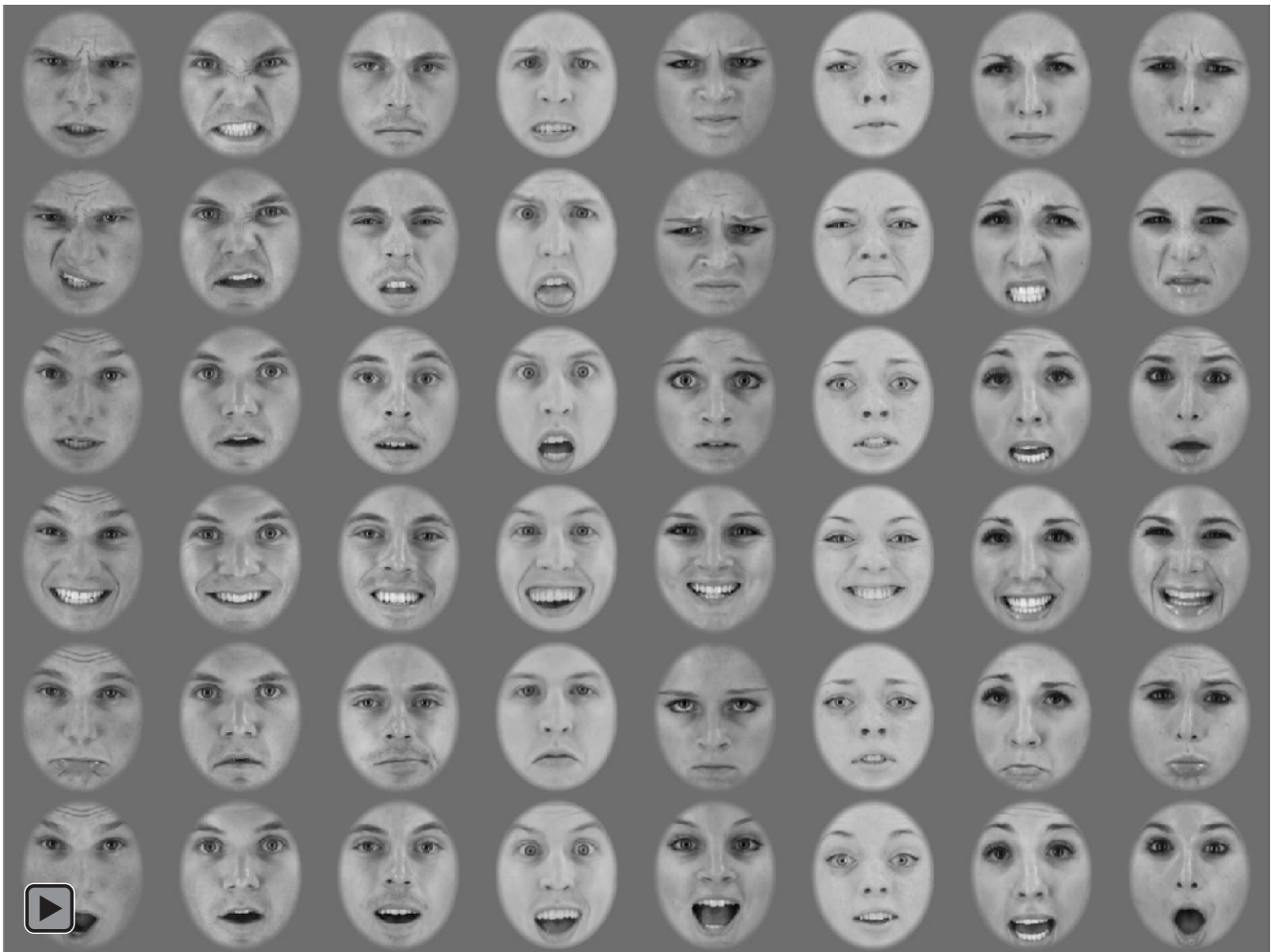


Figure 7. The Reversed movies of each actor (columns) making each facial expression (rows, from top to bottom: anger, disgust, fear, happiness, sadness, surprise).

We tested the idea that the final frames of the expressions carried the most relative information by independently measuring the ideal observer's performance at each frame in the Dynamic movies (Figure 6). We found that indeed, the ideal observer's threshold was highest for the first movie frame (where all actors had a neutral expression—although not identical neutral expressions, as evidenced by the ideal observer's ability to perform the task) and progressively decreased across successive frames. These results offer an interesting new insight into the temporal evolution of information in human facial expressions: Namely, they show that human facial expressions may become systematically more informative as they unfold over time.

Finally, our finding that efficiency did not significantly differ for Dynamic and Shuffled expressions suggests that human observers are remarkably insensitive to the temporal structure of the flow of

information during the production of a facial expression. Although this may seem like an implausible result, Figure 3 demonstrates that it is surprisingly easy to identify an actor's expression when the order of the stimulus frames is randomly permuted. Observers in our experiment shared this subjective impression, which was ultimately borne out in their task performance and information processing efficiency.

Our results with Dynamic and Shuffled expressions are consistent with the idea that observers do not require the contiguous temporal progression of information that takes place during the creation of natural dynamic expressions in order to efficiently extract information from them. If so, this predicts that observers' efficiencies should also not be affected by other temporal transformations, such as the reversal of the expression events. We tested this prediction by measuring thresholds and efficiencies for the same Dynamic and Shuffled stimuli as in our main experi-

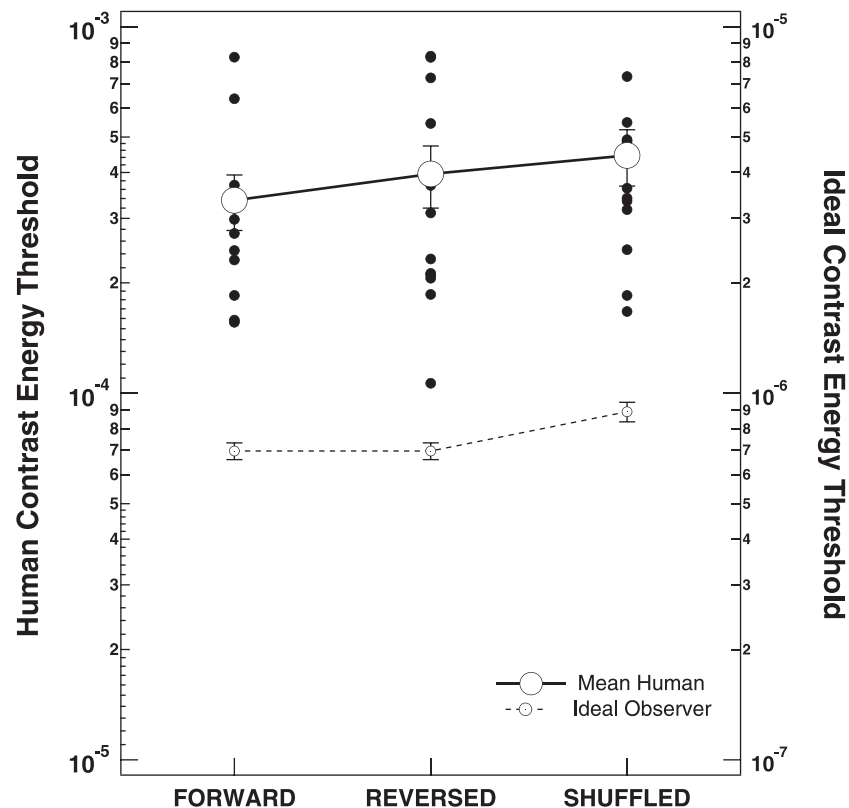


Figure 8. Contrast energy thresholds for human observers (small solid circles, large open circles and solid line, left ordinate) and the ideal observer (small open circles and dashed line, right ordinate) for the Dynamic, Reversed, and Shuffled facial expression stimuli. Each solid circle corresponds to an individual human participant; large open circles correspond to the mean threshold across human participants in each condition. Error bars on each mean correspond to ± 1 SEM.

ment, along with a Reversed condition in which all of the Dynamic expressions were shown in reverse temporal order (Figure 7). Six authors and seven new naive observers participated in the experiment. All other aspects of the stimuli and experimental design were identical to those in the original experiment.

The results of this follow-up experiment are shown in Figures 8 and 9. Figure 8 plots thresholds for the ideal observer (dashed line and small open circles), individual human observers (small closed circles), and the mean across human observers (solid line and large open circles). The data from one naive observer was removed from the analysis due to an inability to perform the task above 40% correct in all conditions. First, note that the ideal observer's thresholds are identical for the Dynamic and Reversed stimuli, because the stimuli are equally informative in both conditions. Second, the pattern across conditions was very similar for our human observers: Mean thresholds for Forward and Reversed expressions were nearly identical and, as in our original experiment, thresholds were higher for the Shuffled expressions. The corresponding efficiencies are plotted in Figure 9. These data show that efficiency was nearly the same across all three conditions. A one-way repeated measures analysis of variance confirmed that

there were no significant differences across conditions, $F(2, 22) = 0.01$, $p = 0.99$.

Conclusions

Establishing how much information is carried by different aspects of facial expressions, as well as how much each contributes to our ability to discern emotion, is of fundamental importance to ultimately achieving a general understanding of how emotions are implicitly communicated. Although our intuition may strongly suggest that the dynamic properties of facial expressions should play an important role in this process, the results of previous experiments designed to evaluate the contribution of dynamic facial cues to the recognition of emotion have been equivocal. A large part of this inconsistency may stem from a general confounding of the physical availability of information with the psychological ability to make use of information. Our experiments have shown that when these two factors are carefully dissociated, the presence of dynamic facial cues appears to offer no additional benefit to emotion recognition beyond what is already

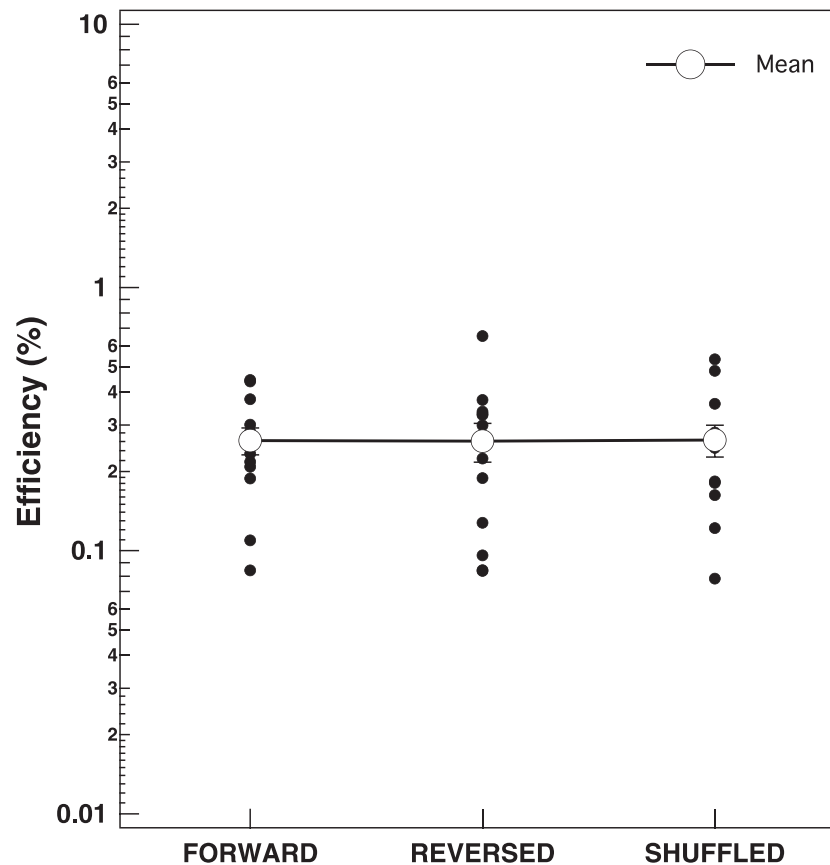


Figure 9. Corresponding efficiencies (ideal/human contrast energy thresholds) for each human participant shown in Figure 8. Error bars on each mean correspond to ± 1 SEM.

provided by a single, static snapshot of a facial expression.

Keywords: ideal observer, face perception, facial expressions, efficiency

Acknowledgments

All authors contributed to the design and execution of the experiments. J. D. B. and J. M. G. generated the stimuli. J. M. G. carried out the analyses and wrote the initial draft of the manuscript. All authors discussed the results and provided extensive comments on subsequent drafts of the manuscript. We would like to thank two anonymous reviewers for helpful comments and for the suggestion to include the follow-up experiment on reversed expressions. This research was supported by National Institute of Health grant R01-EY019265.

Commercial relationships: none.

Corresponding author: Jason M. Gold.

Email: jgold@indiana.edu.

Address: Department of Psychological and Brain

Sciences, Indiana University, Bloomington, Indiana, USA.

References

- Ambadar, Z., Schooler, J. W., & Cohn, J. F. (2005). Deciphering the enigmatic face: The importance of facial dynamics in interpreting subtle facial expressions. *Psychological Science*, *16*(5), 403–410.
- Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality & Social Psychology*, *70*(3), 614–636.
- 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 & Social Psychology*, *37*(11), 2049–2058.
- Bould, E., & Morris, N. (2008). Role of motion signals in recognizing subtle facial expressions of emotion. *British Journal of Psychology*, *99*(Pt 2), 167–189.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*(4), 433–436.

- Bruce, V., Henderson, Z., Greenwood, K., Hancock, P. J. B., Burton, A. M., & Miller, P. (1999). Verification of face identities from images captured on video. *Journal of Experimental Psychology: Applied*, *5*, 339–360.
- 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*(4), 233–240.
- Christie, F., & Bruce, V. (1998). The role of dynamic information in the recognition of unfamiliar faces. *Memory & Cognition*, *26*(4), 780–790.
- Clarke, T. J., Bradshaw, M. F., Field, D. T., Hampson, S. E., & Rose, D. (2005). The perception of emotion from body movement in point-light displays of interpersonal dialogue. *Perception*, *34*(10), 1171–1180.
- 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]
- Edwards, K. (1998). The face of time: Temporal cues in facial expressions of emotion. *Psychological Science*, *9*(4), 270–276.
- Efron, B., & Tibshirani, R. (1993). *An introduction to the bootstrap*. New York: Chapman & Hall.
- Ekman, P. (1972). Universals and cultural differences in facial expressions of emotion. In J. Cole (Ed.), *Nebraska symposium on motivation, 1971* (pp. 207–283). Lincoln: University of Nebraska Press.
- Ekman, P. (1993). Facial expression and emotion. *American Psychologist*, *48*(4), 384–392.
- Fiorentini, C., & Viviani, P. (2011). Is there a dynamic advantage for facial expressions? *Journal of Vision*, *11*(3):17, 1–15, <http://www.journalofvision.org/content/11/3/17>, doi:10.1167/11.3.17. [PubMed] [Article]
- Fujimura, T., & Suzuki, N. (2010). Effects of dynamic information in recognising facial expressions on dimensional and categorical judgments. *Perception*, *39*(4), 543–552.
- Geisler, W. S. (1989). Sequential ideal-observer analysis of visual discriminations. *Psychological Reviews*, *96*(2), 267–314.
- Geisler, W. S. (2004). Ideal observer analysis. In J. S. Werner & L. M. Chalupa (Eds.), *The visual neurosciences* (p. 824–837). Cambridge, MA: MIT Press.
- Geisler, W. S. (2011). Contributions of ideal observer theory to vision research. *Vision Research*, *51*(7), 771–781.
- Gold, J. M., Bennett, P. J., & Sekuler, A. B. (1999a). Identification of band-pass filtered letters and faces by human and ideal observers. *Vision Research*, *39*(21), 3537–3560.
- Gold, J. M., Bennett, P. J., & Sekuler, A. B. (1999b). Signal but not noise changes with perceptual learning. *Nature*, *402*(6758), 176–178.
- Gold, J. M., Tadin, D., Cook, S. C., & Blake, R. (2008). The efficiency of biological motion perception. *Perception & Psychophysics*, *70*(1), 88–95.
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
- Kamachi, M., Bruce, V., Mukaida, S., Gyoba, J., Yoshikawa, S., & Akamatsu, S. (2001). Dynamic properties influence the perception of facial expressions. *Perception*, *30*(7), 875–887.
- Katsyri, J., & Sams, M. (2008). The effect of dynamics on identifying basic emotions from synthetic and natural faces. *International Journal of Human-Computer Studies*, *66*, 233–242.
- Knappmeyer, B., Thornton, I. M., & Bulthoff, H. H. (2003). The use of facial motion and facial form during the processing of identity. *Vision Research*, *43*(18), 1921–1936.
- Lander, K., & Bruce, V. (2000). Recognizing famous faces: Exploring the benefits of facial motion. *Ecological Psychology*, *12*(4), 259–272.
- Lander, K., Christie, F., & Bruce, V. (1999). The role of movement in the recognition of famous faces. *Memory & Cognition*, *27*(6), 974–985.
- Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. Cambridge, UK: Cambridge University Press.
- O'Toole, A. J., Phillips, P. J., Weimer, S., Roark, D. A., Ayyad, J., Barwick, R., et al. (2011). Recognizing people from dynamic and static faces and bodies: Dissecting identity with a fusion approach. *Vision Research*, *51*(1), 74–83.
- Pelli, D. G. (1990). The quantum efficiency of vision. In C. Blakemore (Ed.), *Vision: Coding and efficiency* (pp. 3–24). Cambridge, MA: Cambridge University Press.
- Tjan, B. S., Braje, W. L., Legge, G. E., & Kersten, D. (1995). Human efficiency for recognizing 3-D objects in luminance noise. *Vision Research*, *35*(21), 3053–3069.
- Wehrle, T., Kaiser, S., Schmidt, S., & Scherer, K. R. (2000). Studying the dynamics of emotional expression using synthesized facial muscle movements. *Journal of Personality & Social Psychology*, *78*(1), 105–119.