Categorical Imperative *NOT*: Facial Affect is Perceived *Continuously*

Diane J. Schiano	**Sheryl M. Ehrlich**	**Kyle Sheridan**
CSLI	Adobe Systems-MSE11	MDExpert.com
Stanford University	321 Park Avenue	419 Waverly Street
Stanford CA 94304	San Jose CA 95110	Palo Alto, CA 94301
diane@psych.stanford.edu	sehrlich@adobe.com	Kyle@mdexpert.com
+ 1 650 854 1948	+ 1 408 536 3786	+1 650 327 7999

ABSTRACT

Facial affect (or emotion) recognition is a central issue for many VMC and naturalistic computing applications. Most computational models assume "categorical perception" of facial affect, in which a benign illusion promotes robust recognition of emotional expressions even under severe degradation conditions, including temporal compression. However, this applied interest in human facial affect perception is coming at a time when the evidence for categorical perception is being challenged in the basic research literature, largely on methodological grounds. The research presented here systematically addresses the classic evidence for categorical perception of facial affect, using high-quality digital imaging and display technologies and improved research methods. In doing so, it illustrates a fruitful convergence of basic and applied research. The evidence does *NOT* support categorical perception of facial affect, which in turn suggests the importance of preserving high-fidelity motion information in portraying emotion. This research provides new human behavioral data on facial affect perception, and underscores the importance of careful consideration of facial affect compression methods.

Author Keywords

Affective computing, avatars, facial affect**,** facial expression of emotion, affect, emotion, face, nonverbal communication. video compression, VMC

ACM Classification Keywords

H.5.1 Information Interfaces and Presentation (e.g., HCI). Multimedia Information Systems. Evaluation/methodology.

INTRODUCTION

Of all the visual stimuli we encounter in the world, human faces and facial expressions of emotion are clearly among the most significant and salient. Emotion (or "affect") is central to human experience, and facial displays are our primary means of perceiving and expressing emotion. These facial gestures comprise a natural and compelling medium for human communication, and there is evidence that this may also be true for computer interfaces [4,11, 14,17,18,24]. Such findings, together with recent developments in video image capture, display and recognition--and in broadband data transmission technologies--have produced a surge of interest in facial affect perception in the HCI community.

Researchers working in the fields of video mediated communication (VMC) and naturalistic computer interfaces are especially interested in perception of facial affect. "Affective computing" AI models, based on established human behavioral data, are currently being developed to recognize facial expressions of emotion in video, with many intriguing potential applications (e.g., unobtrusive product testing, stress & lie detection, affective avatars, improved facial recognition models, and much more) [14,17]. The possibilities for the future seem limitless, but current applications are still quite limited, especially in terms of bandwidth and compute-times. Using video in such applications is possible only under highly constrained or compressed conditions. Often the choice is made to show highly realistic facial images (i.e., high spatial resolution, full color) at the expense of motion (e.g., reduced video frame rates, motion compression, time lags). Yet it is becoming increasingly clear that this may not be the appropriate tradeoff [4]. Facial recognition is robust across severe image and resolution degradations [1]. However, perception of emotional content relies heavily on motion information [9]. Emotions can be identified in displays consisting solely of fields of moving dots (i.e., no facial features depicted whatsoever) [2]. The timing and trajectories of facial expressions are highly precise (some are in the millisecond range), and people show great sensitivity to temporal parameters in trying to determine felt emotion, its intensity and sincerity [9]. This suggests that if a bandwidth tradeoff is required, one should consider preserving high-fidelity motion information at the expense of image realism, not the other way around.

CHI 2004, April 24–29, 2004, Vienna, Austria.

Curiously, the growing applied interest in perception of facial affect is coming at a time when much of the established human behavioral data on the topic is being challenged in the basic research literature, largely on methodological grounds. Schiano and colleagues [21,22] discuss these issues at some length, and we have addressed several of these concerns in our previous work. The research described in the present paper focuses on the question of whether perception of facial affect should indeed be considered "categorical" in nature, a basic psychological issue with potentially profound applied impact.

Categorical perception refers to an illusory phenomenal tendency towards assimilation (enhanced similarity) of items within a perceptual category, and contrast (enhanced distinction) between categories. Some theorists hold that categorical perception reflects fundamental features of human cognition and neural functioning [10]; others, that it reflects properties peculiar to only a few special stimulus domains [3,13]. In any case, such effects have proven useful in applications in areas as diverse as speech (phoneme) recognition [13] and color perception [3]. The question of whether facial expressions are perceived as qualitatively distinct categories or varying more continuously along some underlying dimension(s) has been a subject of intense debate in recent psychological research [7,20,21]. That categorical perception effects should be found for facial affect is somewhat surprising, since faces—and facial expressions—clearly vary along multiple physical dimensions [21] (not to mention that perception of emotion may involve many complex psychological, social and situational factors). However, recent findings supporting the categorical perception view (many by one group of researchers) have been highly influential in the applied community [5,16,25]. Thus, current computational models of facial affect perception now routinely characterize it as categorical in nature [6,14,16].

Categorical perception is classically assessed through two defining paradigms, specific types of identification and discrimination tasks. Prototypical exemplars from two distinct categories are morphed to create an interpolated continuum of equal physical intervals. Typically, participants are asked to respond to each stimulus using a "two-alternative forced-choice" method, with the continuum endpoints as the response alternatives. Categorical perception predicts that identification should ideally be at asymptote for category prototypes at the endpoints of each continuum, with a sharp boundary region of abrupt change in the middle values. As illustrated in Figure 1, this is essentially the convergence of two (out of phase) step functions The Figure illustrates the ideal (hypothetical) results indicating categorical perception in identifying stimuli transitioning between the emotional expressions happy and sad.

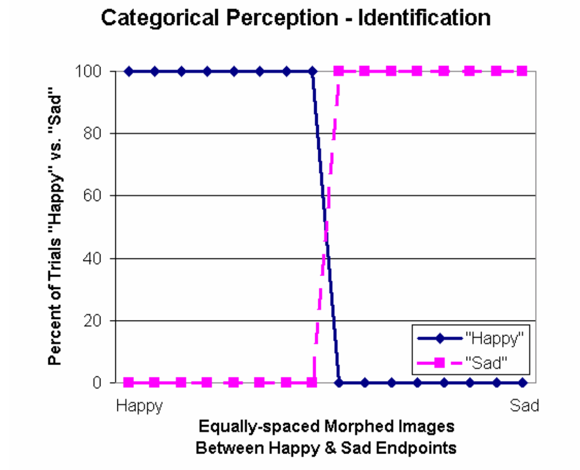

Figure 1. Illustration of ideal (hypothetical) results indicating categorical perception in identifying stimuli transitioning between Happy and Sad expressions.

In a related vein, categorical perception predicts that the ability to discriminate between equally-spaced stimuli transitioning between two category endpoints should peak in the middle values in the category boundary region. A "same-different" discrimination task is typically used. The ideal (hypothetical) results for discriminating stimuli transitioning between the emotional expressions happy and sad are shown in Figure 2.

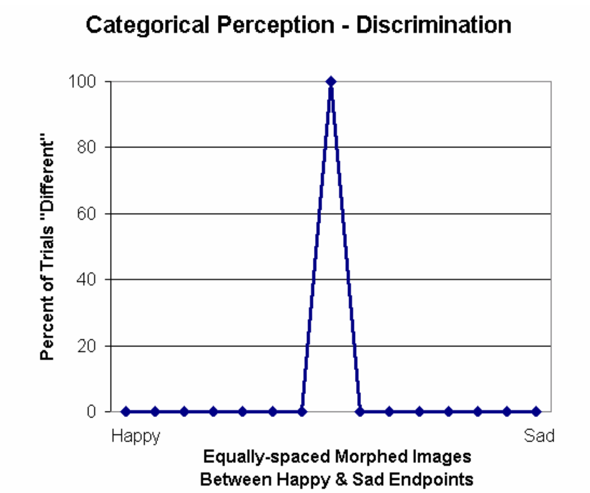

Figure 2 Illustration of ideal (hypothetical) results indicating categorical perception in discriminating stimuli transitioning between Happy and Sad expressions.

Most human facial expressions have an extended temporal trajectory that may involve several seconds, but they can also change abruptly--sometimes in a matter of milliseconds. The net effect of categorical perception of facial affect would be to blur the ability to distinguish between successive views of a given emotion during its trajectory, but also to sharpen the detection of a distinct new emotion once it has occurred. That is, an emotional expression should be perceived fairly consistently across its trajectory, while changes in emotion should still be quickly detected, even under substantially degraded conditions. Such an effect--if it is found to hold--could be very useful for affective applications, promoting robust perception of emotions even under substantial temporal compression. However, recent criticisms of previous research methods now make those findings suspect [19,20,21]. The primary purpose of the present research was to see if the evidence still supports categorical perception of facial affect when the studies are performed using the same classic experimental paradigms, but improved materials and methods.

Previous studies in support of categorical perception of facial affect suffer from a host of methodological flaws. First, almost all past work on facial expression of emotions uses stimuli derived from a single source. This is Ekman & Friesen's [8] corpus of black-and-white still photographs of a few highly-trained actors portraying 6 theoretically "basic" emotional expressions: Anger, disgust, fear, happiness, sadness, surprise. The stimuli were often low-quality, line-drawings or fairly low-resolution monochrome computer images derived from these photographs. Morphing software was used to construct stimulus continua transitioning between pairs of these independent images, rather than from actual physical transitions. Moreover, few studies included baseline or "neutral" (no emotion) stimuli. Secondly, strong criticisms have been made that the standard two-alternative (i.e., continuum endpoint) “forced-choice" response format used in the classic identification task could by its very nature distort results to make continuous perception appear categorical [19,20]. Finally, the standard same-different discrimination task was typically implemented *sequentially* (the "ABX" procedure), due to image display constraints. However, this imposed an unnecessary memory load, which could bias results to reflect contrastive cognitive differentiation effects. This could easily yield results apparently similar to, but actually quite different from, true perceptual categorization. More detailed critiques of previous research, and more complete discussion of the approaches we devised to overcome these limitations, are available elsewhere [21,22].

In the studies to be presented here, we examine the evidence for categorical perception of facial affect using the standard experimental paradigms (the identification and discrimination tasks), with improved materials and measures. Extremely high quality (high-resolution, full color) stimuli were created from real-time videos of actors physically transitioning between emotional expressions, including "neutral", or no expression. Stimulus images consisted of frames sampled from these images ("naturalistic" condition) or of morphed transitions between video end-frames ("morphed" condition). The validity of these stimuli was confirmed in earlier work [21,22]. Experiment 1 involves an identification task across each emotion pair continuum, using an "open-ended multiple-choice + rating” response format, rather than the standard two-alternative forced-choice format, to limit bias. A comparison between identification patterns for morphed and natural transition conditions is included. Experiment 2 uses a *simultaneous* same-different discrimination task, to minimize memory load. Pairs of images chosen from equally-spaced intervals across each continuum of morphed emotion pairs (with some falling within, and others between, category boundaries), were used as stimuli. Both discrimination accuracy and reaction-time data were recorded. Since the data derived from these studies is more complex than those of previous research, many more--and more complex--analyses were required to thoroughly assess the evidence. Due to space constraints, the results of only a few primary analyses can be presented and discussed in the pages to follow. However, further findings will be made available upon request.

EXPERIMENT 1: IDENTIFICATION TASK

Experiment 1 uses the classic identification task for testing categorical perception. In this experiment, participants are asked to identify the emotion expressed in a series of stimulus images. The images were either directly sampled from naturalistic (i.e., physical) transitions made between two emotions, or from morphs made at equally- weighted intervals between the endpoint images. To avoid potential bias, the open-ended multiple-choice + rating response format was used, rather than the standard two-alternative forced-choice format. With this format, participants are shown a list of all the emotional expressions (including neutral) depicted in the set of images, and are asked to rate the extent to which each emotion is seen in a given image. Multiple responses and "other" (open-ended) responses are permitted. The categorical perception predictions for this identification task are essentially the same as those illustrated in Figure 1, since simply adding response options should not affect results stemming from truly *categorical* perception. Thus, clear evidence for categorical perception of facial affect would be found if identification of a given category is close to either at 0%o or 100% at either end of each continuum, and remains there until a very sharp boundary region of abrupt change occurs in the middle values. To the extent or that multiple responses are given, that identification rates do not remain at asymptote until a region of abrupt change, or that the boundary region shows gradual rather than abrupt change, the evidence would not support categorical perception of facial affect. This would suggest that facial expressions of emotion--even under highly simplified conditions--are perceived in a much more complex way than the categorical perception hypothesis suggests. For simplicity of reference, we use the term "continuous perception" for any pattern of findings in this research that is clearly not categorical in nature.

Methods

Participants

40 UC Berkeley and Stanford University students ranging from 18 to 30 years of age participated for pay.

Materials

High-resolution color images of facial expressions made by 2 drama students (1 female, 1 male) from San Francisco State University produced the facial expression stimuli use in this study. (These actors' images were chosen empirically from a larger library we had created, because they could be morphed with minimal "ghosting" and other extraneous feature distortions). The actors were briefly shown some of Ekman's standard images for each emotion during an initial orientation session. To promote naturalness, the actors were then instructed to simply imagine a time when they strongly felt each emotion (the recommended approach [7,12]). Each actor provided a total of 14 different front-view exemplars of the six basic emotions (anger, disgust, fear, happiness, sadness, surprise). The actors were then asked to transition between all pairwise combinations of Ekman's "six basic emotions" (anger, disgust, fear, happiness, sadness, surprise) plus a "neutral" expression (no expression). This yielded 42 transitions (e.g., from happy to sad, happy to neutral, sad to happy) per actor and 7 "identity" clips in which the actor held one expression for the entire clip. The duration of each clip was eight seconds. The recorded transitions were then digitized at full resolution, 30 frames per second, onto a Macintosh PowerPC using a Targa DTX digital video card. In these experiments we used full-color stills taken from the high-resolution video images. We used both morphs (created with Gryphon Morph on the Macintosh) and real transitions (to maintain realistic, possible facial muscle movements rather than linear interpolations of the images). Sample frames from the natural and morphed angry to happy transitions for one actor are shown below in Figure 3.

Previous work drew upon the Ekman stills; the present studies used our database of facial expressions posed by 2 actors. The accuracy of depiction was confirmed in our previous research [21,22]. Stimuli were presented using a HyperCard stack, designed for this experiment, in color on a 20" computer monitor connected to a Macintosh PowerPC.

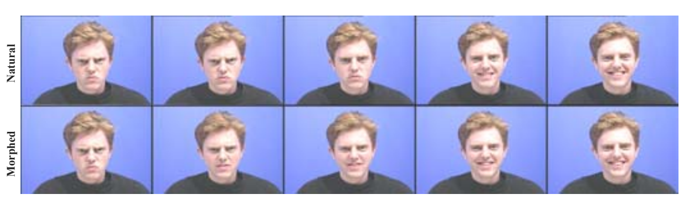

Figure 3. Sample images from the naturalistic and morphed Angry to Happy transitions for one actor.

Procedure

Participants viewed still frames taken at 1-second intervals from the digitized video clips and morphs of all pairwise combinations of the six basic emotions plus neutral of one of two actors, including the first and last frames. A separate set of participants viewed stimuli from a second actor. A list of the six emotions (anger, disgust, fear, happiness, sadness, surprise) plus "neutral" was displayed on the computer monitor concurrently as each still frame was displayed. Participants rated the degree to which each of the listed emotions was present in each of the stimuli on a scale of 0 (not present at all) to 6 (extremely high). Participants also had the option to type in any emotion they thought was displayed by a stimulus that was not listed and to rate that term. A practice session preceded testing in which participants viewed 10 randomly selected frames not viewed during test trials. Participants then viewed each test frame once, for a total of 216 test trials. Ratings for each emotion were recorded on every trial by the HyperCard program. The stimuli were presented in random order. Participants proceeded at their own pace and were encouraged to take breaks as needed. The entire procedure took about an hour.

Results and Discussion

Previous categorical perception of facial affect studies have used the classic two-alternative forced-choice response format which requires the participant to choose which one of two endpoint emotions a given stimulus depicts. Using this approach, it is fairly straightforward to plot the proportion of times on which the participant selected each endpoint. However, since we used the open-ended multiple choice + ratings response format, our data are much more complex, requiring many more--and more complex--analyses to understand. As discussed earlier, we can only discuss a small subset of the analyses performed here, but additional information is available upon request.

In the categorical perception literature, the results of the identification task are graphed in terms of identification rates for two category endpoints and the continuum between them, as illustrated in Figure 1. The results sections of many papers consist almost solely of such graphs for all stimulus pairs. In order to make our results directly comparable to those previous findings, we first collapsed our dataset to yield a single (i.e., the highest-rated) response for each trial. We then determined all "other" emotions, across the continuum for each transition pair. We did this for both the naturalistic and morphed transitions, averaged over all relevant stimuli from both actors. This process, while informative, yielded too many graphs to present in a single paper. In Figure 4, we show a subset of the resulting curves for transitions beginning with anger and ending in either fear, happiness or sadness. These were chosen as a conservative test of the categorical perception predictions, since we had previous evidence suggesting that of all emotions, anger displays may be most likely to be perceived categorically [22]. Results for both naturalistic and morphed transitions are shown in the figure.

Even a cursory look at these sets of curves suggests some difficulties for categorical perception predictions. The fact that identification rates were generally high at either end of the continua is not surprising, since the stimuli were chosen to be good exemplars of each emotion (with the possible exception of fear, which is discussed in some detail elsewhere [21,22]).

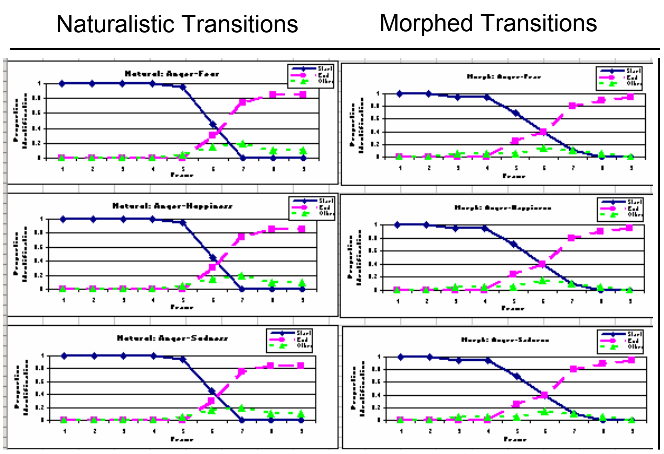

Figure 4. Identification functions for highest-rated response for naturalistic & morphed transitions between Anger and either Fear, Happiness or Sadness in Experiment 1.

Moreover, the curves had to cross at some point. The primary question here is whether this crossover was done abruptly or gradually. As can be seen in the graphs, the results varied for the different emotional expressions, but overall, the boundary regions are clearly not as uniformly abrupt as predicted by categorical perception. Moreover, even under conditions in which only the single highest-rated response was plotted, and even for a clearly identified emotion such as anger, participants responded with "other" emotions fairly often, especially in the boundary regions, but even towards the endpoints (especially for fear, which is notoriously easily confused with other emotions--most notably surprise, as we found in this and previous studies). These findings do not gel well with categorical perception predictions.

In the first set of analyses, we plotted the results for only a single response per stimulus. However, participants in this study were able to indicate multiple responses, if they felt this was appropriate. Categorical perception predicts that only a single response will be given, regardless of response options. We therefore performed another set of analyses, looking at the proportion of transitions in which participants gave multiple responses. Two different criteria were used to provide an estimate of multiple responding, depending on the number of transition frames (out of a total of 9) for which multiple responses were given. Using the strict criterion, if a multiple response were given to any 1 of the 9 stimuli across the continuum, the transition was counted as demonstrating multiple responses. Using the lenient criterion, only those transitions for which multiple responses were given for 3 of the 9 images were counted as demonstrating multiple responses. This analysis was performed for both naturalistic and morphed transitions.

Figure 5 is a summary graph showing the overall proportion of transitions for which multiple responses were given for both natural and morphed stimuli. A strict interpretation of categorical perception would argue that multiple responses should not be made at all, or should be limited to an abrupt boundary region, perhaps 1 of the 9 images per continuum. However, the incidence of multiple responding is surprisingly high, and similar for both types of stimuli. That is, even using the lenient criterion, over 60% of the transitions--for both natural and morphed stimuli--were given multiple responses. Using the strict criterion, this rose to almost 90%. These data are strong evidence against categorical perception, since *more than half* of the transitions showed multiple responses in *a third or more* of the images in each continuum (for which the endpoints were considered good exemplars of each emotion).

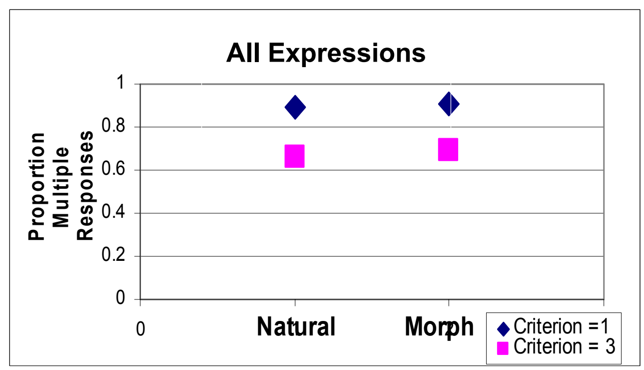

Figure 5. Proportion of naturalistic and morphed transitions given multiple responses, using two criterion conditions.

In an attempt to map our findings better onto the previous research, a further analysis considered only multiple responses from the endpoints of each continuum (i.e., a trial was considered as showing a multiple response only if both endpoints of a given continuum were given as responses). The results are shown in Figure 6. Even under these highly constrained conditions, a high proportion of multiple responding was found. Naturalistic and morphed transitions again show similar patterns of results, although overall the morphed transitions clearly show a lower proportion of multiple responses. This is interesting given that the morphs were created by physically averaging the images, while the naturalistic transitions were performed by the actors. This difference suggests that the categorical results found in past work might be partly due to the exclusive use of morphs and/or the two-alternative forced-choice response format, which permitted only endpoint-expression responses. Interestingly, the morphed transition/endpoint-only condition, most similar to past studies, yields results closest to categorical perception.

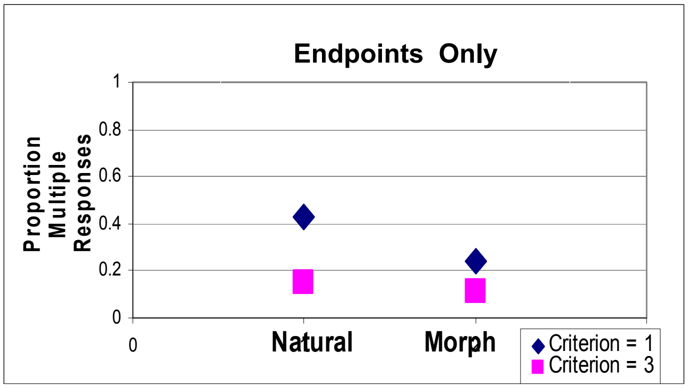

Figure 6. Proportion of naturalistic and morphed transitions given multiple responses from both endpoint expressions only.

EXPERIMENT 2: DISCRIMINATION TASK

Experiment 2 employs the second classic approach to studying categorical perception, the discrimination task. In a discrimination study, participants are shown pairs of images taken at equal intervals across each continuum of emotion pairs. Results are recorded in terms of discrimination accuracy (proportion correct) and reaction-time (to decide whether or not the stimuli are different). Categorical perception predicts that stimulus differences within a perceptual category are not discerned as well as those between categories (i.e., in the boundary region). Thus, if emotional expressions are perceived categorically, discrimination should generally be low, with an abrupt peak in performance (i.e., increased accuracy and perhaps lower reaction times) in the middle values of each continuum, at the category boundary. The predicted curve for discrimination accuracy is illustrated in Figure 2.

Methods

Participants
Twenty Stanford University students ranging from 18 to 30 years of age participated in this study for pay.

Materials
Stimuli were a subset of those of the first experiment. Once again, we used full-color stills taken from the high-resolution video images. However, we only tested the morphed transitions because one cannot assume equally-spaced transitions (necessary for this argument) in natural changes in emotion. We further restricted the test set of transitions to include only transitions between angry, fearful, happy, and sad for one actor due to time constraints. Stimuli were presented using a HyperCard stack, designed for this experiment, in color on a 20" video monitor connected to a Macintosh PowerPC.

Procedure
Participants viewed pairs of still frames sampled from the morphed transition clips taken at two-step intervals (frames 1+3, 2+4, …). The task was to indicate whether the pairs were the same or different by making a keyboard press on keys marked "S" and "D". The use of a simultaneous A-B same-different method (rather than an ABX paradigm) eliminated any possible memory effects. A practice session preceded testing in which participants viewed 10 comparison pairs not viewed during test trials. Participants viewed a total of 1080 test trials presented in random order; each of the 2-step pairs (of interest here) was viewed 5 times. Same-different responses and reaction times for each pair were recorded on every trial by the HyperCard program. Participants proceeded at their own pace, and were encouraged to take breaks as needed. The entire procedure took about 1-1.5 hours for most participants.

Results and Discussion

If facial expressions of emotion are perceived categorically, we would expect a peak in discrimination between equally-spaced at the border between categories. Thus, although steps between presented stimuli from the morphs were physically equal, discrimination accuracy should be fairly low and then abruptly peak in the middle values, at the boundary region. Figure 7 plots discrimination accuracy (in terms of proportion correct) across all participants for each stimulus transition pair. In the figure, each line represents a difference emotion pair continuum (see key for specific emotion pairs); each data point represents 20 trials.

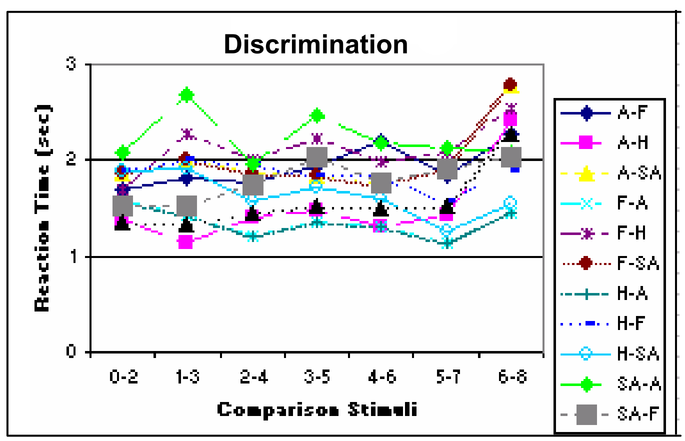

Figure 7. Proportion correct discrimination for 2-step stimulus images across each transition between emotion pairs.

Even under cursory inspection, these results clearly do not support categorical perception predictions. Discrimination performance varies for the different stimulus pairs, but in general show no clear peak in performance. Indeed, the performance curves are generally flat, especially across the middle values (with the one exception of anger-fear, which shows a gently rising curve).

Another way to look at discrimination task performance is in terms of reaction-time to decide whether images are the same or different. Figure 8 plots discrimination reaction-time for correct responses across all participants for each stimulus transition pair. As in Figure 7, each line in Figure 8 represents a difference emotion pair continuum (see key for specific emotion pairs) and each data point represents 20 trials. Categorical perception of facial affect predicts fairly high reaction-times at the continuum endpoints (i.e., within perceptual categories), with a sudden "dip" (or inverse peak) in the middle values, in the boundary region.

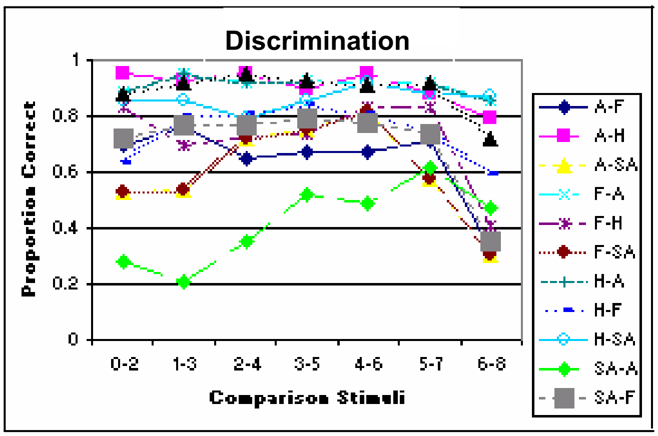

Figure 8. Reaction time for correctly discriminating 2-step stimulus images across each transition between emotion pairs.

Again, this is *not* what we found.

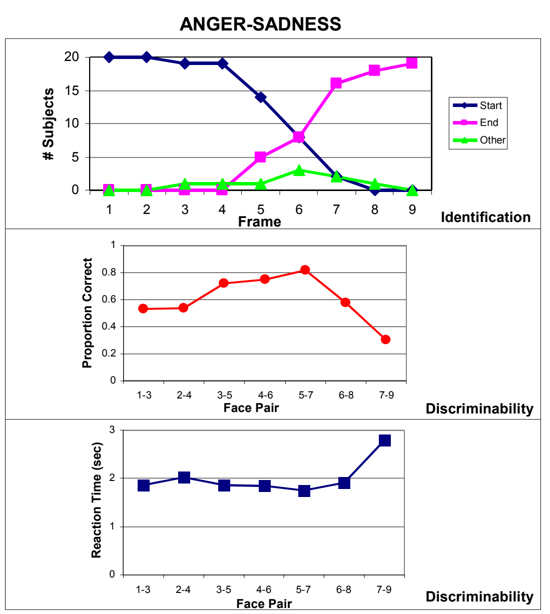

Figure 9. Compiled identification and discrimination data across the Anger-Sadness morphed transitions.

Instead, the reaction-time curves, while somewhat variable, are fairly flat overall, and especially in the middle values; no "dip" is apparent. These reaction-time data converge well with the correct discrimination data. Thus, the results of Experiment 2--both in terms of discrimination accuracy and reaction-time for correct discrimination--show no evidence in support of categorical perception of facial affect. Indeed, the discrimination evidence argues strongly against categorical perception predictions, even for specific transitions pairs where the identification data may not be as clear, as seen in Figure 9.

GENERAL DISCUSSION

Categorical perception of facial affect is a complex and controversial research topic. While most researchers find the evidence for categorical perception compelling for at least a few salient and simple stimulus domains [3,13], there are those who see it as the foundation of human cognition in general [10], and others who regard it as a myth [15]. Most agree, however, that much of the previous research on perception of facial affect in general--and categorical perception of facial affect in particular--suffers from a variety of serious methodological flaws. The research described here used the same experimental paradigms as previous research, but improved materials and methods. The evidence does *not* support categorical perception of facial affect, even under these highly simplified research conditions.

The studies presented here are part of a larger research program focused on characterizing the perception of facial expressions of emotion and exploring the potential applications of our findings. The present results converge well with our previous work establishing new approached for measuring and modeling perception of facial affect [21], and exploring the role of (real-time video) motion cues in disambiguating degraded (compressed) images of facial expressions of emotion [9]. They complete the work-in-progress we reported on earlier [23], and provide the first set of comparable data for morphed and naturalistic transitions between facial expressions.

Affective computing models are critically dependent upon the human behavioral data from which they are derived. It is hoped that that the data provided in this research will be used to help re-examine the assumption of categorical perception of facial affect in current computational models. In addition, designers of video-mediated communication and affective interface applications require accurate human behavioral data to inform the tradeoffs they must make in dealing with limited bandwidth and long compute times. Of course, bandwidth demand has always exceeded supply and we expect it to continue to do so, even in the broadband future. The critical issue is how to use the bandwidth that is available most effectively. Our data suggest that decisions affecting visual dynamics (video frame rate, compression algorithms, time lags,) must be approached with extreme caution to avoid timing biases which can dramatically affect perceived affect, and hence the nature of the communication [4]. The hope that the "benign illusion" of categorical perception would serve to provide special protection for facial affect perception against such degradations itself appears illusory.

REFERENCES

1. Bachmann, T. Identification of spatially quantized tachistoscope images of faces. *European Journal of Cognitive Psychology, 3* (1991), 87-103.

2. Bassili, J.N. Emotion Recognition. *Jnl of Personality and Social Psychology, 37* (1979), 2049-2058.

3. Bornstein, M.H. & Korda, N.O. (1984). Discrimination and matching within and between hues measured by reaction times. Some implications for categorical perception. *Psychological Research, 46*, 207-222

4. Bruce, V. The role of the face in communication. *Interacting with Computers, 8,* (1996), 166-176.

5. Calder, A., Young, A., Perrett, D., Etcoff, N.,Rowland, D. (1996). Categorical perception of morphed facial expressions. *Visual Cognition,* (3), 81-117.

6. Dailey, M.N., Cottrell, G.W., Padgett, C., & Adolphs, R. (2002), EMPATH: A neural network that categorizes facial expressions. *Jnl Cognitive Neuroscience*, *14* (8), 1158 1173.

7. Ekman. P. (Ed.) (1982). *Emotion in the human face*. Cambridge: Cambridge University Press.

8. Ekman, P. & Friesen, W.V. (1978). *The Facial Action Coding System.* Palo Alto, Consulting Psychologists.

9. Ehrlich, S.M., Schiano, D.J. & Sheridan, K. (2000). Communicating facial affect: It's not the realism; it's the motion. *Proc. CHI' 2000*, 252-253. NY: ACM.

10. Harnad, S. (2003) Categorical Perception,in *Encyclopedia of Cognitive Science*. NY: MacMillan.

11. Kurlander, D., Skelly, T. & Salesin, D. (1996). Comic chat. *Proc. Computer Graphics*, ACM Press, 225-236.

12. Levenson, R. W., Ekman, P., Friesen, W. V. Voluntary facial action generates emotion-specific autonomic nervous system activity. *Psychophysiology, 27* (1990), 363-384.

13. Liberman, A., Harris, K., Hoffman, H & Griffith, B. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Jnl of Experimental Psychology*, *54*, 358-368.

14. Lisetti, C.L. & Schiano, D.J. (2000). Automatic facial expression interpretation: Where Human-Computer Interaction, Artificial Intelligence & Cognitive Science intersect. *Pragmatics and Cognition*, *8* (1), 185-235.

15. Massaro, D.W. (1998). Categorical perception: Important phenomenon or lasting myth? *Proc. ICSLP'98*. Sydney, Austrlia.

16. Padgett, C. & Cottrell, G. W. (1998). A simple neural network models categorical perception of facial expressions. *Cognitive Science*, 806-807.

17. Picard, R. (1997). *Affective computing*. Cambridge: MIT Press.

18. Reeves, B. & Nass, C.I. (1996). *The media equation: How people treat computers, television, & new media like real people and places.* NY: Cambridge Univ. Press.

19. Russell, J.A. Is there universal recognition of emotion from facial expression? *Psychological Bulletin, 95* (1994), 102-141.

20. Russell, J.A., & Fernandez-Dols, J.M. (1997). *The psychology of facial expression*. NY: Cambridge Univ. Press.

21. Schiano, D.J., Ehrlich, S.M., Rahardja, K. & Sheridan, K. (2000). Measuring and modeling facial affect. *Behavior Research Methods, Instruments & Computers, 32* (4), 505-514.

22. Schiano, D.J., Ehrlich, S.M., Rahardja, K. & Sheridan, K. (2000). Face to Interface: Facial Affect in(Hu)Man and Machine. *Proc. CHI'2000*, 193-200. NY: ACM.

23. Schiano, D.J., Ehrlich, & Sheridan, K. (2001). Categorical Perception of Facial Affect: An Illusion. *Ext. Abstracts, CHI'2001*, 299-300. NY: ACM.

24. Walker, J.H., Sproull, L., Subramani, R. (1994). Using a human face in an interface. *Proc. CHI'1994*, 205. Boston, MA.

25. Young, A., Rowland, D., Calder, A., Etcoff, N., Seth, A. Perrett, D. (1997). Facial expression megamix. *Cognition*, 63, 271-31.