12-1-2006

Prior Experience as a Stimulus Category Confound: an Example using Facial Expressions of Emotion

Leah H. Somerville
Dartmouth College

Paul J. Whalen
Dartmouth College

Tools of the Trade

Prior experience as a stimulus category confound: an example using facial expressions of emotion

Leah H. Somerville and Paul J. Whalen

Department of Psychological & Brain Sciences, Center for Cognitive Neuroscience, Dartmouth College, Hanover, NH, USA

Facial expressions of emotion represent a stimulus set widely used to assess a broad range of psychological processes. However, a consideration of systematic differences between expression categories, other than differences relating to characteristics of the expressions themselves, has remained largely unaddressed. By collecting experience rankings in a large sample of undergraduates, we observed that the amount of reported experience individuals have had with different facial expressions of emotion systematically differed between all expression categories. These findings shed light on the potential for identifying confounds inherent to comparing some stimulus categories and, in this case, may aid in the interpretation of observed between-expression category findings.

Humans comprise an especially challenging subject pool precisely because we as experimenters lack the ability to control their prior experiences, a form of control that is more readily employed in animal studies. As Tolman put it, 'rats live in cages; they do not go on binges the night before one has planned the experiment' (Tolman, 1945, p. 166). Cognitive neuroscientists have devised clever experimental manipulations that have offered important insights into the neural substrates of everything from memory and attention to mood and moral reasoning. But given the experientially 'tainted' nature of our subjects of study, careful consideration should be given to whether prior experience differs systematically between any two experimental comparison conditions. To this end, the current report summarizes and discusses one such instance from the realm of affective neuroscience, namely, differential prior experience with primary emotional facial expression categories.

Images depicting facial expressions of emotion comprise standardized stimulus sets that are frequently employed to assess a wide range of psychological functions (e.g. Ekman and Friesen, 1976; Russell, 1994). While one can question the ecological validity of presenting subjects with static 2D images of expressions, a vast amount of useful and replicable data have been gleaned from these stimuli. However, a consideration of systematic differences between expression categories, other than differences relating to characteristics of the expressions themselves, has remained largely unaddressed.

Psychologists studying emotion have long pointed out that as important as fear states are, they are relatively rare events (Cannon, 1927). William James summarized the issue as follows: 'The progress from brute to man is characterized by nothing so much as by the decrease in frequency of proper occasions for fear' (James, 1890, p. 415). If it is true that we experience the feeling of fear less often than other emotions, then it follows that we would see fearful expressions on the faces of others less often than other expressions over the course of our lifetime (Bond and Siddle, 1996; Whalen, 1998). Such a familiarity bias would be an important caveat to consider when interpreting dependent measurements associated with responses to fearful facial expressions in comparison to other more frequently encountered expressions. Accordingly, the current report characterizes the past experience individuals report having with distinct primary expression categories in their lifetimes.

METHODS

Subjects were 1537 undergraduate students tested in groups of ~200. All procedures were conducted in accordance with the Institutional Review Board of the University of Wisconsin. After excluding improperly ranked surveys, 1393 usable subjects remained, consisting of 774 females and 619 males. Subjects were presented with a list of the six primary facial expression labels (Ekman and Friesen, 1976) and the neutral expression label, and were instructed to rank order these seven expression labels based upon the frequency

with which they believed they had encountered these expressions in their lifetimes (1 = seen the most, 7 = seen the least). To increase clarity, two words were presented together to describe each expression category.

RESULTS

Label, mean ranking and standard deviation (in parentheses) of rankings for each expression label were as follows: happy/smiling: 1.86 (1.26); neutral/expressionless: 2.27 (1.75); sad/unhappy: 3.65 (1.25); angry/mad: 4.03 (1.29); surprised/startled: 4.89 (1.52); disgusted/yuck: 5.37 (1.40) and fearful/afraid: 5.93 (1.33). A Friedman rank test for related samples yielded a significant main effect for expression, indicating that the distributions of ranks differed across expression $X_F^2(6) = 4218.26$, $P < 0.001$, $r = 0.87$. These ranking distributions are presented as a frequency histogram by expression in Figure 1. Given the significant omnibus effect, Wilcoxon signed rank tests were performed on each pair of neighboring expressions to determine whether the distributions of ranks differed significantly from one another. Reported results are corrected for multiple comparisons, and effect size (r) is denoted after significance values. The distributions of all neighboring pairs differed significantly from one another: happy/neutral: $Z = 4.92$, $P < 0.0001$, $r = 0.132$; neutral/sad: $Z = 19.15$, $P < 0.0001$, $r = 0.51$; sad/angry: $Z = 7.85$, $P < 0.0001$, $r = 0.21$; angry/surprise: $Z = 13.28$, $P < 0.0001$, $r = 0.36$; surprise/disgust: $Z = 8.09$, $P < 0.0001$, $r = 0.22$; disgust/fear: $Z = 85.45$, $P < 0.0001$, $r = 0.29$. Separate analyses of male and female subjects produced the same ranked order of expressions.

DISCUSSION

Here we define the pattern of reported differential prior experience with distinct facial expressions of emotion in a large sample of undergraduates. Rank position was significantly different between all adjacently ranked expression categories, with fearful expressions consistently reported as having been seen the least in one's lifetime, and happy expressions being seen the most. These data offer a caution for behavioral and neuroimaging studies seeking to compare responses to one expression with that of another, and more generally to any experimental paradigm involving conditions

where stimuli fundamentally differ on the dimension of prior experience.

Here we consider the specific example of amygdala reactivity to human faces. In human neuroimaging studies, fearful facial expressions evoke greater responsivity in the amygdala compared to neutral or happy expressions (Morris et al., 1996; Whalen et al., 1998), consistent with the notion that the amygdala responds to fearful faces more because they have predicted negative outcomes in the past. The present report raises the possibility that greater amygdala reactivity to fearful faces might also reflect a heightened response to an environmental stimulus that has been encountered less frequently in one's lifetime. Such a notion would be consistent with experimental findings showing that the amygdala is more responsive to novel compared to familiar faces (DuBois et al., 1999; Schwartz et al., 2003; Wright et al., 2003).

While the present results have specific implications for interpreting data that vary across facial expression categories, they imply that for any measured difference between facial categories, one should consider the possible confound of prior life experience. For example, greater amygdala activity has been reported in Caucasian-American subjects while viewing unfamiliar African-American faces than unfamiliar Caucasian-American faces (Hart et al., 2000; Phelps et al., 2000; Cunningham et al., 2004), which could be attributed to the relative novelty of African-American faces to Caucasian-American subjects (but see Lieberman et al., 2005). Supporting this possibility is a study by Phelps et al. (2000) that attempted to equate novelty of different race faces by using only pictures of highly familiar, famous Caucasian-American and African-Americans, and found comparable amygdala activity to the two stimulus conditions (Phelps et al., 2000). Further, while amygdala activation to African-American faces has been shown to vary with scores on the Implicit Association Task (IAT; Greenwald et al., 1998), a purported measure of implicit racism, it is possible to interpret IAT effects in terms of familiarity, where reaction time biases track categories that have been encountered more frequently (Rothermund and Wentura, 2004; Kinoshita and Peek-O'Leary, 2005).

Within the specific example of facial expression categories, some experimental designs have eliminated this alternative explanation by utilizing paradigms that manipulated other
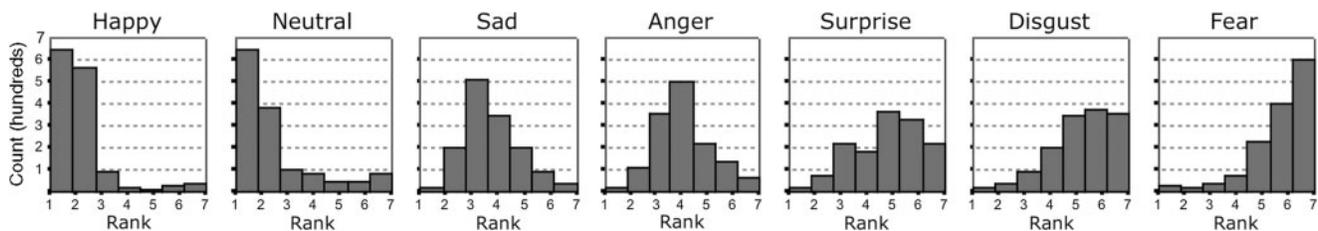
Fig. 1 Distributions of experience rankings with seven facial expressions of emotion. Distributions are ordered from left to right according to frequency ranking (1 = seen the most, 7 = seen the least). Y-axis denotes frequency.

aspects of the face presentation, such as eye gaze (Adams and Kleck, 2003), while holding facial expression itself constant. For example, a neuroimaging study by Kim *et al.* (2004) used a paradigm employing only surprised faces while manipulating the valence of contextual information describing these faces. The resulting neural activity observed reflected differences in the interpretation of valence associated with these expressions (as a function of context) and could not be explained by the differences in experience reported here (see also Adams *et al.*, 2003).

In addition, differential prior experience with facial expression categories would likely influence recognition difficulty. For example, labeling accuracy for the facial expressions depicted in the Pictures of Facial Affect (Ekman and Friesen, 1976), a commonly used facial expression stimulus set, is best for happy expressions (95–98% agreement) and worst for fearful expressions (77–87% agreement) (Ekman and Friesen, 1976; Russell, 1994). Although cross-cultural studies indicate that accuracy decreases when judging emotions of faces in other cultures, happiness and fear are still the most and the least accurately identified, respectively (for a review, see Elfenbein and Ambady, 2002). These data resonate with expression labelling accuracy rates observed in brain-damaged patients, in that negative expressions are consistently labelled less accurately than positive expressions (Braun *et al.*, 2005; Broks *et al.*, 1998; Rapcsak *et al.*, 2000; Rosen *et al.*, 2006).

It should be noted that these data rely on subjects' ability to accurately quantify this facet of past experience. These findings would be supplemented by additional studies assessing the incidence of observed expressions across a number of real-life social settings (see Bond and Siddle, 1996 for preliminary data that address this issue). In addition, it would be useful to determine whether the same ranked order would be found using different eliciting stimuli (e.g. different category labels, using faces rather than verbal labels as items to rank), and other formats of evaluation such as use of a Likert scale or free response format. Indeed, a free response format (rather than the forced choice format) might reveal other experienced expression categories not considered here, such as social emotions like embarrassment (Keltner and Buswell, 1997).

In the future, the employment of this scale could provide insight into the psychological state of the individual, or group, of interest. An individual's rank ordering, if abnormal with respect to an overall sample, could be related to differences in one's past experiences and/or be predictive of certain personality traits or tendencies toward psychopathology. For example, a small minority of our sample rated sad faces as being seen the most. Future studies could aim to show that such a response relates to symptoms of depression, since prior research has demonstrated a tendency for depressed individuals to over-ascribe sad emotion to faces (Gur *et al.*, 1992). A similarly small number of subjects rated angry faces as being seen the most.

Such a response could be shown to be related to experiences during one's upbringing, given that individuals raised in an abusive or neglectful environment ascribe anger in a presented face at a lower threshold (Pollak and Sinha, 2002).

Numerous studies have demonstrated robust differences in behavioral and psychophysiological response to these primary facial expression categories. The present demonstration that these categories differ based upon reported past experience further complicates the interpretation of observed between-category findings. Our hope is that further research on this issue will lead to an improved understanding of these ubiquitously employed stimuli, as well as our processing of these primary expression categories. More generally, these data suggest that differential past experience will be an important variable to consider when comparing different stimulus categories within cognitive neuroscience research.

Conflict of Interest

None declared.

REFERENCES

Adams, R.B., Gordon, H.L., Baird, A.A., Ambady, N., Kleck, R.E. (2003). Effects of gaze on amygdala sensitivity to anger and fear faces. *Science*, *300*, 1536.

Adams, R.B., Kleck, R.E. (2003). Perceived gaze direction and the processing of facial displays of emotion. *Psychological Science*, *14*, 644–7.

Bond, N.W., Siddle, D.A.T. (1996). The preparedness account of social phobia:Some data and alternative explanations. In: Rapee, R.M., editor. *Current Controversies in the Anxiety Disorders,* New York: Guilford Press, pp. 291–316.

Braun, M., Traue, H.C., Frisch, S., Deighton, R.M., Kessler, H. (2005). Emotion recognition in stroke patients with left and right hemispheric lesion: Results with a new instrument – the FEEL test. *Brain and Cognition*, *58*, 193–201.

Broks, P., Young, A.W., Maratos, E.J., et al. (1998). Face processing impairments after encephalitis: Amygdala damage and recognition of fear. *Neuropsychologia*, *36*, 59–70.

Cannon, W.B. (1927). The James-Lange theory of emotion: A critical examination and an alternative theory. *American Journal of Psychology*, *39*, 10–124.

Cunningham, W.A., Johnson, M.K., Raye, C.L., Gatenby, C.J., Gore, J.C., Banaji, M.R. (2004). Separable neural components in the processing of black and white faces. *Psychological Science*, *15*, 806–13.

DuBois, S., Rossion, B., Schlitz, C., et al. (1999). Effect of familiarity on the processing of human faces. *Neuroimage*, *9*, 278–89.

Ekman, P., Friesen, W.V. (1976). *Pictures of Facial Affect*. Palo Alto, CA: Consulting Psychologists Press.

Elfenbein, H.A., Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin*, *128*, 203–35.

Greenwald, A.G., McGhee, D.E., Schwartz, J.K.L. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology*, *74*, 1464–80.

Gur, R.C., Erwin, R.J., Gur, R.E., Zwil, A.S., Heimberg, C., Kraemer, H.C. (1992). Facial emotion discrimination: II. Behavioral findings from depression. *Psychiatry Research*, *42*, 241–51.

Hart, A.J., Whalen, P.J., Shin, L.M., McInerney, S.C., Fisher, H., Rauch, S.L. (2000). Differential response in the human amygdala to racial outgroup vs ingroup face stimuli. *Neuroreport*, *11*, 2351–5.

James, W. (1890). *The Principles of Psychology, Vol. II*, New York: Henry Holt.

Keltner, D., Buswell, B.N. (1997). Embarrassment: Its distinct form and appeasement functions. *Psychological Bulletin*, 122, 250–70.

Kim, H., Somerville, L.H., Johnstone, T., et al. (2004). Contextual modulation of fMRI responsivity to surprised faces. *Journal of Cognitive Neuroscience*, 16, 1730–45.

Kinoshita, S., Peek-O'Leary, M. (2005). Does the compatibility effect in the race Implicit Association Test reflect familiarity or affect? *Psychonomic Bulletin & Review*, 12, 442–52.

Lieberman, M.D., Hariri, A., Jarcho, J.M., Eisenberger, N.I., Bookheimer, S.Y. (2005). An fMRI investigation of race-related amygdala activity in African-American and Caucasian-American individuals. *Nature Neuroscience*, 8, 720–2.

Morris, J.S., Frith, C.D., Perrett, D.I., et al. (1996). A differential neural response in the human amygdala to fearful and happy facial expressions. *Nature*, 383, 812–5.

Phelps, E.A., O'Connor, K.J., Cunningham, W.A., et al. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. *Journal of Cognitive Neuroscience*, 125, 729–38.

Pollak, S.D., Sinha, P. (2002). Effects of early experience on children's recognition of facial displays of emotion. *Developmental Psychology*, 38, 784–91.

Rapcsak, S.Z., Galper, S.R., Comer, J.F., et al. (2000). Fear recognitions after focal brain damage: A cautionary note. *Neurology*, 54, 575–81.

Rosen, H.J., Wilson, M.R., Schauer, G.F., et al. (2006). Neuroanatomical correlates of impaired recognition of emotion in dementia. *Neuropsychologia*, 44, 365–73.

Rothermund, K., Wentura, D. (2004). Underlying processes in the Implicit Association Test (IAT): Dissociating effects of salience and valence. *Journal of Experimental Psychology: General*, 133, 139–65.

Russell, J.A. (1994). Is there a universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychological Bulletin*, 118, 379–91.

Schwartz, C.E., Wright, C.I., Shin, L.M., et al. (2003). Differential amygdalar response to novel versus newly familiar neutral faces: a functional MRI probe developed for studying inhibited temperament. *Biological Psychiatry*, 53, 854–62.

Tolman, E.C. (1945). A stimulus-expectancy need-cathexis psychology. *Science*, 101, 160–6.

Whalen, P.J. (1998). Fear, vigilance, and ambiguity: Initial neuroimaging studies of the human amygdala. *Current Directions in Psychological Science*, 7, 177–88.

Whalen, P.J., Rauch, S.I., Etcoff, N.L., McInerney, S.C., Lee, M.B., Jenike, M.A. (1998). Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *Journal of Neuroscience*, 18, 411–8.

Wright, C.I., Martis, B., Schwartz, C.E., Shin, L.M., Fisher, H.H., McMullin, K. (2003). Novelty responses and differential effects of order in the amygdala, substantia innominata, and inferior temporal cortex. *Neuroimage*, 18, 660–9.