

Emotion

Valence Resolution of Ambiguous Facial Expressions Using an Emotional Oddball Task

Maital Neta, F. Caroline Davis, and Paul J. Whalen

Online First Publication, June 27, 2011. doi: 10.1037/a0022993

CITATION

Neta, M., Davis, F. C., & Whalen, P. J. (2011, June 27). Valence Resolution of Ambiguous Facial Expressions Using an Emotional Oddball Task. *Emotion*. Advance online publication. doi: 10.1037/a0022993

Valence Resolution of Ambiguous Facial Expressions Using an Emotional Oddball Task

Maital Neta, F. Caroline Davis, and Paul J. Whalen
Dartmouth College

Previous research suggests that neural and behavioral responses to surprised faces are modulated by explicit contexts (e.g., "He just found \$500"). Here, we examined the effect of implicit contexts (i.e., valence of other frequently presented faces) on both valence ratings and ability to detect surprised faces (i.e., the infrequent target). In Experiment 1, we demonstrate that participants interpret surprised faces more positively when they are presented within a context of happy faces, as compared to a context of angry faces. In Experiments 2 and 3, we used the oddball paradigm to evaluate the effects of clearly valenced facial expressions (i.e., happy and angry) on default valence interpretations of surprised faces. We offer evidence that the default interpretation of surprise is negative, as participants were faster to detect surprised faces when presented within a happy context (Exp. 2). Finally, we kept the valence of the contexts constant (i.e., surprised faces) and showed that participants were faster to detect happy than angry faces (Exp. 3). Together, these experiments demonstrate the utility of the oddball paradigm to explore the default valence interpretation of presented facial expressions, particularly the ambiguously valenced facial expression of surprise.

Keywords: oddball, facial expressions, ambiguity, emotion

Supplementary material: <http://dx.doi.org/10.1037/a0022993.sup>

Facial expressions provide information about the emotions and intentions of others. One basic signal communicated by facial expressions is the valence of the emotion being experienced by the expressor and, in turn, the valence of the outcome predicted for the perceiver. For example, happy and angry faces are consistently rated as positive and negative, respectively (Ekman & Friesen, 1976). However, ratings of other facial expressions suggest that their inherent valence is more ambiguous. For example, surprised facial expressions have signaled both positive and negative events in the past, and in the absence of contextual information that can be used to disambiguate the valence of this expression, some people interpret surprised faces negatively, while others interpret them positively (Kim et al., 2003; Neta et al., 2009; Neta & Whalen, 2010). In this way, surprised expressions offer a means to assess individual differences in one's positivity-negativity bias.

This dual-valence representation for surprise makes this facial expression especially dependent on the context in which it is encountered (see Bouton, 1994 for a similar argument made for

stimuli encountered during extinction training). Consistent with this notion, clearly valenced "contextual" stimuli have been shown to modulate ratings of ambiguously valenced "target" facial expressions (Kim et al., 2004; Russell & Fehr, 1987). For example, *explicit* verbal contexts—that is, a context that specifically defines the valence of a facial expression by offering an explanation for what triggered it, such as "He just lost \$500" or "She just won the lottery," modulated valence ratings of surprised faces in a negative or positive direction, respectively (Kim et al., 2004). Similarly, judgments of neutral faces (i.e., targets) are influenced by the presence of other clearly valenced facial expressions (i.e., happy, sad) presented as a context (Russell & Fehr, 1987). Thus, even an *implicit* context—that is, a context that does not offer a specific explanation, but rather implies only whether the environment is positively or negatively valenced, can influence the interpretation of ambiguously valenced facial expressions.

In Experiment 1, we sought to extend these implicit contextual modulation effects to the facial expression of surprise. Note that the Kim et al. (2004) contextual modulation of surprise study, mentioned above, showed a priming effect where explicit verbal contexts primed subjects to interpret the valence of surprised faces in a way that was congruent with the contextual information. Conversely, Russell and Fehr (1987) observed a contrast effect such that neutral faces were interpreted more negatively when presented in the context of happy faces, and more positively when presented within the context of sad faces (the so-called "anchoring effect"). These effects were observed whether the other faces were presented simultaneously with the target faces or sequentially.

The differing outcomes observed by Russell and Fehr (1987) and Kim et al. (2004) could be related to at least two method-

Maital Neta, F. Caroline Davis, and Paul J. Whalen, Department of Psychological and Brain Sciences, Dartmouth College.

We thank G. Wolford for statistical advising and comments on the manuscript, and C. J. Norris for advice on experimental design. We also thank R. A. Loucks and M. H. Miner for assistance in recruitment of participants and data collection, and E. J. Ruberry and A. L. Duffy for input on experiment design. This work was supported by NIMH 080716.

Correspondence concerning this article should be addressed to Maital Neta, Department of Psychological & Brain Sciences, Dartmouth College, 6207 Moore Hall, Hanover, NH 03755. E-mail: maital.neta@dartmouth.edu

ological differences: (a) the use of neutral versus surprised faces as the ambiguous target face and/or (b) the use of explicit (i.e., verbal) versus implicit (i.e., other faces) contextual information. One important difference between surprised and neutral expressions is that, while neutral faces are intended to be devoid of emotional meaning, surprised faces can predict both positive and negative outcomes. That is, when we encounter surprised faces in our social world, the meaning of the expression is defined by the context. As such, we predicted that an implicit context consisting of clearly valenced facial expressions would cause surprised faces to be interpreted as congruent with the context, in contrast to the anchoring effect observed for neutral faces by Russell and Fehr (1987). To this end, Experiment 1 assessed valence ratings of surprised faces when these faces were presented in a context consisting of either happy (positive context) or angry (negative context) faces. We predicted that surprised faces would be rated more positively when presented in the context of happy faces, and more negatively when presented in the context of angry faces.

In previous work, we have shown that surprised expressions can be used to measure individual differences in one's positivity-negativity bias, since there are individual differences in one's propensity to rate surprised expressions as negative or positive (Kim et al., 2003; Neta et al., 2009). We have hypothesized that the more automatic interpretation of surprised faces is initially negative, suggesting that positive interpretations require a greater degree of regulation (Kim et al., 2003). To test this, we filtered these faces to low (LSF) and high spatial frequency (HSF) representations (Neta & Whalen, 2010). Previous data show that LSF visual information is processed in an elemental fashion followed by a finer analysis of HSF information (Bar et al., 2006; Hughes et al., 1996; Kveraga et al., 2007). Results showed that LSF representations of surprised faces are rated more negatively than HSF representations, indicating that the negative interpretation is "first and fast" (i.e., default negativity) when resolving the valence ambiguity of the surprised expression. In Experiments 2 and 3, we designed a version of the classic "oddball" paradigm with ambiguously and clearly valenced emotional expressions as stimuli, in order to provide further evidence for the primacy of negative interpretations of surprise. The logic of the oddball paradigm is that reaction times will be facilitated to the degree that the intended target is indeed an "oddball" (Campanella et al., 2002). Specifically, in Experiment 2, we predicted that if the default interpretation of surprise is negative, these faces would be detected faster in a context of happy faces compared to angry faces. Finally, in Experiment 3, we extended this idea while holding the valence of the context constant (i.e., surprised faces), and measured reaction time for detecting happy and angry faces. Again, if the default interpretation of surprise is negative, we predicted that participants would be faster to identify the happy faces, because they are the "oddball" stimuli in a surprise face context.

Experiment 1

Method

Participants. Thirty-four healthy Dartmouth undergraduates (22 female; ages 18–21, mean age = 18.9) volunteered to participate. All participants had normal or corrected-to-normal vision,

no psychoactive medication, and no significant neurological or psychiatric history. None were aware of the purpose of the experiment, and they were all compensated for their participation through monetary payment or course credit. Written informed consent was obtained from each participant before the session, and all procedures were approved by Dartmouth College Committee for the Protection of Human Subjects. Two participants were excluded for having BDI scores that were above our cutoff for normal adults (score of above 13). As a result, the final sample contained 32 participants (20 females). All included subjects tested within normal limits for depression [Beck Depression Inventory (BDI; Beck et al., 1961; $M \pm SE = 4.81 \pm .67$)]; and anxiety [State-Trait Anxiety Inventory (Spielberger et al., 1988; STAI_S = 36.4 ± 2.06 ; STAI_T = 38.7 ± 1.86)].

Stimuli and Experimental Design

Calculating a bias score. In the first run, three expressions (angry, happy, surprised) of 18 identities from the NimStim face set (nine female, nine male) were presented (Tottenham et al., 2009). Participants were asked to rate the valence of these faces as positive or negative, based on a gut reaction. We used ratings of surprised faces to measure their positivity-negativity bias by calculating the percentage of negative ratings out of all valence ratings for surprise.

Experimental trials. Subsequently, participants were then presented with four runs, each with a block of 100 stimuli, during which they were again asked to rate the valence of each face as positive or negative. For these runs, three expressions (angry, happy, surprised) of six identities from the Ekman face set (three female, three male) were used (Ekman & Friesen, 1976). Half of the blocks consisted of 84 angry faces, and 16 surprised faces (surprised/angry), and the other half of the blocks, consisted of 84 happy faces, and 16 surprised faces (surprised/happy). These blocks alternated throughout the course of the experiment. The screen background was black. Before starting the task, subjects had to fixate on a small white fixation cross in the center of the screen. Faces were presented for 500 ms, and a white fixation was displayed on a black screen for an intertrial interval of 1000 ms (see Figure 1). Ordering of angry and happy blocks, and the response buttons used were counterbalanced across participants. Reaction times and valence ratings were measured by customized experimental control software, and all trials were analyzed.

Following each testing session, participants also completed the Beck Depression Inventory (BDI; Beck et al., 1961), and State-Trait Anxiety Inventory (STAI_S, STAI_T; Spielberger, et al., 1988).

Results

Valence ratings. An expression (angry, happy, surprised/angry, surprised/happy) repeated measures ANOVA of the experimental trials showed a significant main effect of expression ($F(3, 29) = 487.62, p < .001, \eta^2 = .86$). Fisher's LSD pairwise comparisons revealed that angry faces were rated as more negative than surprised faces, which were rated as more negative than happy faces (all p 's $< .001$). Moreover, surprised/angry faces were rated as more negative than surprised/happy faces ($p < .001$; Figure 2a). Importantly, we calculated bias scores (i.e., a bias to interpret surprise positively or negatively, given no contextual

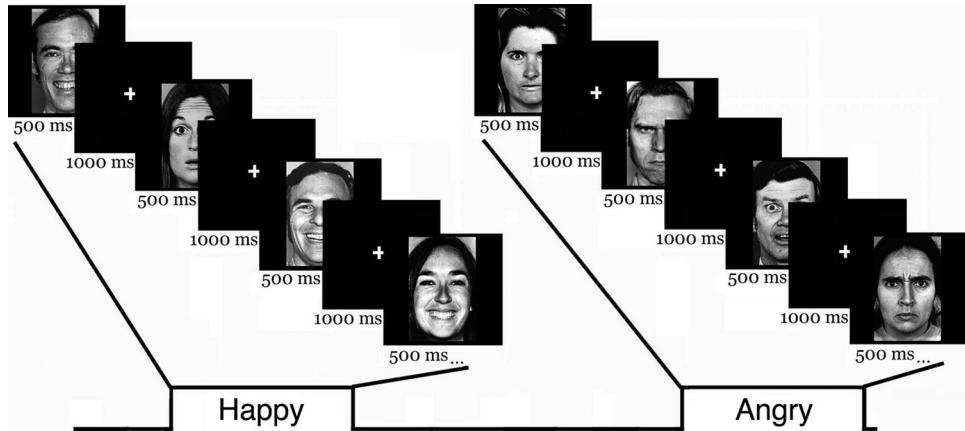


Figure 1. A depiction of the experimental design. Blocks of angry and happy faces, where surprised faces served as oddballs (presented on 16% of the trials in a block). Faces were presented for 500 ms, with a 1000 ms intertrial interval, during which a fixation cross appeared. The task for each face was to decide whether the expression was positive or negative (Exp. 1), or to press one button for each angry/happy face, and a second button each time the expression switched to surprise (Exp. 2). Face images from *Pictures of Facial Affect* by P. Ekman and W. Friesen, 1976, Palo Alto, CA: Consulting Psychologists Press. Copyright 1976 by Consulting Psychologists Press. Reproduced with permission.

information) from each participant in the first run of the experiment, as represented by the dashed line in Figure 2a (percent negative ratings: $M \pm SE = 62.9\% \pm 2.66$). Thus, we compared ratings of surprised faces in the experimental trials to the bias score collected at the start of the experiment and found that participants rated surprise faces more negatively in the angry context, $t(31) = 3.99, p < .001, d_z = .77$, and more positively in the happy context, $t(31) = 2.53, p < .02, d_z = .55$, as compared to their bias. Finally, we also examined a difference score for ratings of surprise within each context versus the ratings for that contextual expression (i.e., ratings of surprised/angry faces minus ratings of angry faces, ratings of surprised/happy faces minus ratings of happy faces). We found that the difference between ratings of surprised/happy and happy faces was greater than the difference between ratings of surprised/angry and angry faces, $t(31) = 3.77, p < .001, d_z = 1.08$.

Reaction times (RTs). An expression (angry, happy, surprised/angry, surprised/happy) repeated measures ANOVA showed a significant main effect of expression ($F(3, 29) = 22.53, p < .001, \eta^2 = .27$). Fisher's LSD pairwise comparisons revealed that participants took longer to rate surprised faces, as compared to angry and happy faces ($p < .001$, except happy vs. surprised/angry = .007). Importantly, there was no difference in RTs for angry and happy faces ($p > .6$) and no difference in RTs between the two conditions of surprised faces ($p > .8$; Figure 2b).

Discussion

The present study showed that, similar to an explicit verbal context (Kim et al., 2004), an implicit context of clearly valenced facial expressions is also sufficient to modulate valence interpretations of surprised faces in a congruent fashion. These results observed to surprised faces are opposite those observed previously to neutral faces, where an implicit context of other facial expressions causes an anchoring (i.e., contrast) in the interpretation of neutral faces. Unlike neutral faces, surprised faces carry a dual-

valence representation that must be resolved, apparently in a fashion suggesting that surprise is better interpreted as congruent with a given context.

Figure 2a also shows that percent negative ratings of surprise within angry blocks were closer to ratings of angry faces than ratings of surprise within happy blocks were to happy faces. In other words, the negative context had a greater influence on biasing ratings of surprise in a negative direction than the positive context did in a positive direction. These data are consistent with (a) our previous report showing that a greater number of subjects tend to lean more heavily in a negative compared to positive direction when interpreting surprise (Neta et al., 2009), and (b) our greater working hypothesis that surprised faces are more inherently negative than positive. To further address this possibility, Experiment 2 modified the classic "oddball" paradigm to present surprised faces in either a clear negative or positive context. If surprised faces are inherently more negative than positive, they will be interpreted as more of an oddball in the positive context (compared to the negative context) and reaction times for detection will be facilitated in the former versus latter context.

Experiment 2

Method

Participants. A new group of 30 healthy Dartmouth undergraduates (13 female; ages 18–22, mean age = 20) volunteered to participate. All participants had normal or corrected-to-normal vision, no psychoactive medication, and no significant neurological or psychiatric history. None were aware of the purpose of the experiment, and they were all compensated for their participation through monetary payment or course credit. Written informed consent was obtained from each participant before the session, and

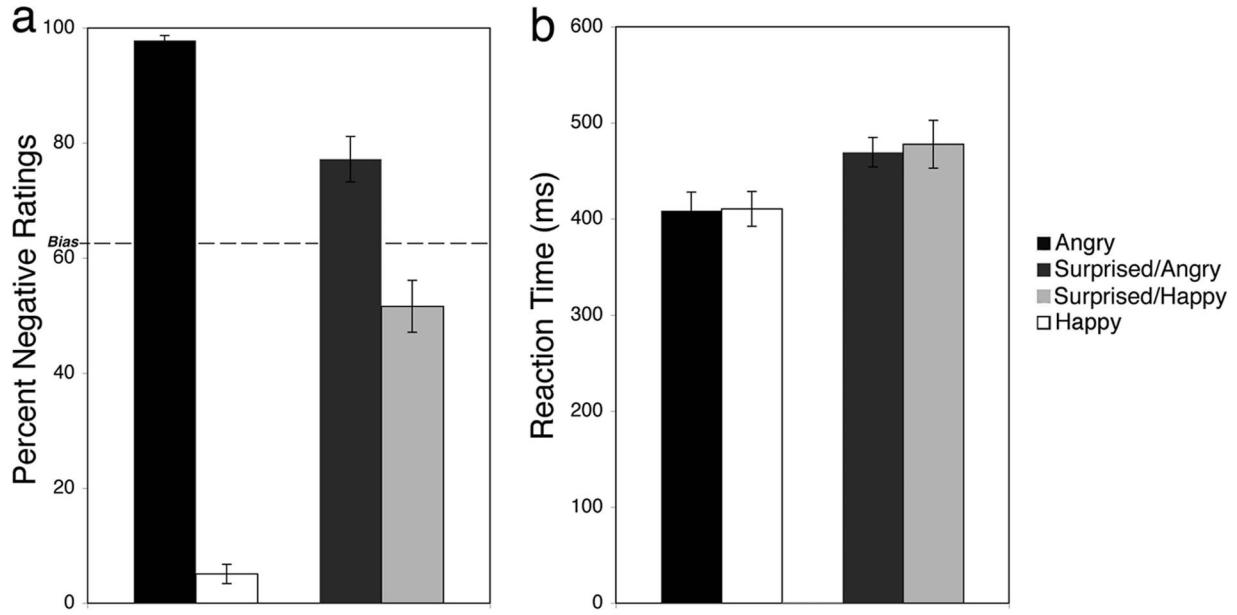


Figure 2. Implicit contexts of clearly valenced facial expressions modulate valence interpretations of surprised faces in a congruent fashion. (a) Surprised faces are rated as more positive in blocks of happy faces, and as more negative in blocks of angry faces, as compared to the mean bias score across participants (represented by the dashed line). (b) There was no difference in reaction times between angry and happy faces, and between surprised faces presented in each context.

all procedures were approved by Dartmouth College Committee for the Protection of Human Subjects. Two participants were excluded for having BDI scores that were above our cutoff for normal adults (score of above 13). As a result, the final sample contained 28 participants (11 females). All included subjects tested within normal limits for depression [Beck Depression Inventory (BDI; Beck et al., 1961; $M \pm SE = 6.42 \pm 1.48$)]; and anxiety [State–Trait Anxiety Inventory (Spielberger et al., 1988; STAI_S = 34.0 ± 2.34 ; STAI_T = 33.0 ± 1.90)].

Stimuli and Experimental Design

The stimuli and procedure in Experiment 2 were identical to those of Experiment 1 with two exceptions. After the first experimental run, participants were asked to perform a new task: to press a button with the right index finger when they detected a surprised face (target) and to press another button with the right middle finger for each angry or happy face (standard). Again, ordering of angry and happy blocks, and the response buttons used were counterbalanced across participants. This task consisted of six total runs. Each run consisted of 100 stimuli, during which a sequence of angry faces or happy faces ('standards,' presented in alternating runs) comprised 84% of trials. Targets consisted of surprised faces, for a total of 16 per run. Reaction times were measured by customized experimental control software, and only trials with correct responses were analyzed.

Results

Reaction times (RTs). An expression (angry, happy, surprised/angry, surprised/happy) repeated measures ANOVA

showed a significant main effect of expression ($F(3, 25) = 65.32$, $p < .001$, $\eta^2 = .76$). Fisher's LSD pairwise comparisons revealed that participants took longer to detect surprised faces in angry blocks, as compared to surprised faces in happy blocks ($p = .001$; Figure 3a). Moreover, they took longer to detect surprised faces, as compared to angry and happy faces, and longer for angry, as compared to happy faces (all p 's $< .001$; $M \pm SE$: angry = 435.6 ± 16.2 , happy = 403.7 ± 12.1 , surprised/angry = 515.6 ± 18.5 , surprised/happy = 485.9 ± 13.3).

Accuracy. An expression (angry, happy, surprised/angry, surprised/happy) repeated measures ANOVA showed a significant main effect of expression ($F(3, 25) = 20.20$, $p < .001$, $\eta^2 = .61$). Fisher's LSD pairwise comparisons revealed that accuracy was higher when detecting angry and happy faces, as compared to surprised faces (p 's $< .001$), but that there was no difference in accuracy for angry and happy faces ($p = .43$) and no difference in accuracy between the two conditions of surprised faces ($p = .78$; $M \pm SE$: angry = $98.5\% \pm .3$, happy = $98.9\% \pm .4$, surprised/angry = $84.9\% \pm 1.9$, surprised/happy = $85.3\% \pm 2.0$; Figure 3b).

Discussion

Experiment 1 provided initial data consistent with the assertion that default valence interpretations of surprised faces are negative, since the percent negative ratings of surprise within angry blocks were closer to ratings of angry faces than ratings of surprise within happy blocks were to happy faces. Experiment 2 provided additional evidence in support of this hypothesis. Surprised faces, given their default negativity, were detected faster in the happy compared to angry context,

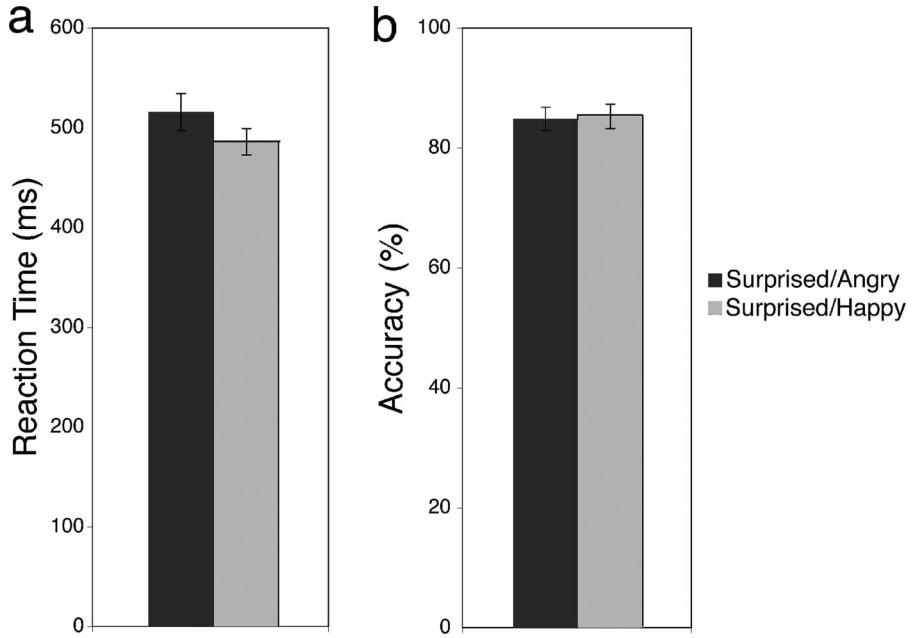


Figure 3. Initial negativity of surprised faces. (a) Surprised faces, given their default negativity, were detected faster in the happy compared to angry context, presumably because the inherently negative surprised expressions stood out as more of an “oddball” in the happy context. (b) There was no difference in accuracy between angry and happy faces (not shown), and between surprised faces presented in each context.

presumably because the inherently negative surprised expressions stood out as more of an “oddball” in the happy context. Experiment 3 was designed to provide additional support for this hypothesis by keeping the context constant (i.e., surprised faces) and then measuring reaction times to angry or happy faces as the oddball stimuli. If surprised faces are indeed inherently negative, reaction times to happy faces will be faster than reaction times to angry faces—because happy represents more of an oddball in the context of surprised faces.

Experiment 3

Method

Participants. A new group of 12 healthy Dartmouth undergraduates (8 female; ages 18–40, mean age = 19.9) volunteered to participate. All participants had normal or corrected-to-normal vision, no psychoactive medication, and no significant neurological or psychiatric history. None were aware of the purpose of the experiment, and they were all compensated for their participation through monetary payment or course credit. Written informed consent was obtained from each participant before the session, and all procedures were approved by Dartmouth College Committee for the Protection of Human Subjects. All subjects tested within normal limits for depression [Beck Depression Inventory (BDI; Beck, Ward, & Mendelson, 1961; $M \pm SE = 2.58 \pm .76$)]; and anxiety [State–Trait Anxiety Inventory (Spielberger, Gorusch, & Lushene, 1988; STAIs = 29.3 ± 2.48 ; STAIr = 31 ± 2.25)].

Stimuli and Experimental Design

The stimuli and procedure in Experiment 3 were identical to those of Experiment 2 with one exception: the target and standard

expressions were flipped, such that each block consisted of 84 trials of surprised faces, and the 16 target trials were an equal mix of angry and happy faces (eight angry and eight happy faces per block), in a randomized order. Again, reaction times were only analyzed for trials with correct responses.

Results

Reaction times (RTs). An expression (angry, happy, surprised) repeated measures ANOVA showed a significant main effect of expression ($F(2, 10) = 18.97, p < .001, \eta^2 = .64$). Fisher’s LSD pairwise comparisons revealed that participants took longer to detect angry faces, as compared to happy faces ($p = .016$). Moreover, they took longer to detect angry ($p = .007$) and happy ($p < .001$), as compared to surprised faces (Figure 4a).

Accuracy. An expression (angry, happy, surprised) repeated measures ANOVA showed a significant main effect of expression ($F(2, 10) = 5.43, p = .025, \eta^2 = .32$). Fisher’s LSD pairwise comparisons revealed that accuracy was higher when detecting happy, as compared to angry faces ($p = .027$), but that there was no difference in accuracy for angry and happy, as compared to surprised faces (p ’s $> .05$; $M \pm SE$: angry = $75.9\% \pm 5.7$, happy = $83.9\% \pm 4.6$, surprised = $92.0\% \pm 3.5$; Figure 4b).

Discussion

Experiment 3 provides additional support for the results in Experiment 2 by using surprised faces as the contextual stimuli and then measuring reaction times to angry or happy faces as the oddball stimuli. Indeed, in the context of surprised faces, reaction times to happy faces were faster than reaction times to angry faces, suggesting that happy faces represent more of an oddball because

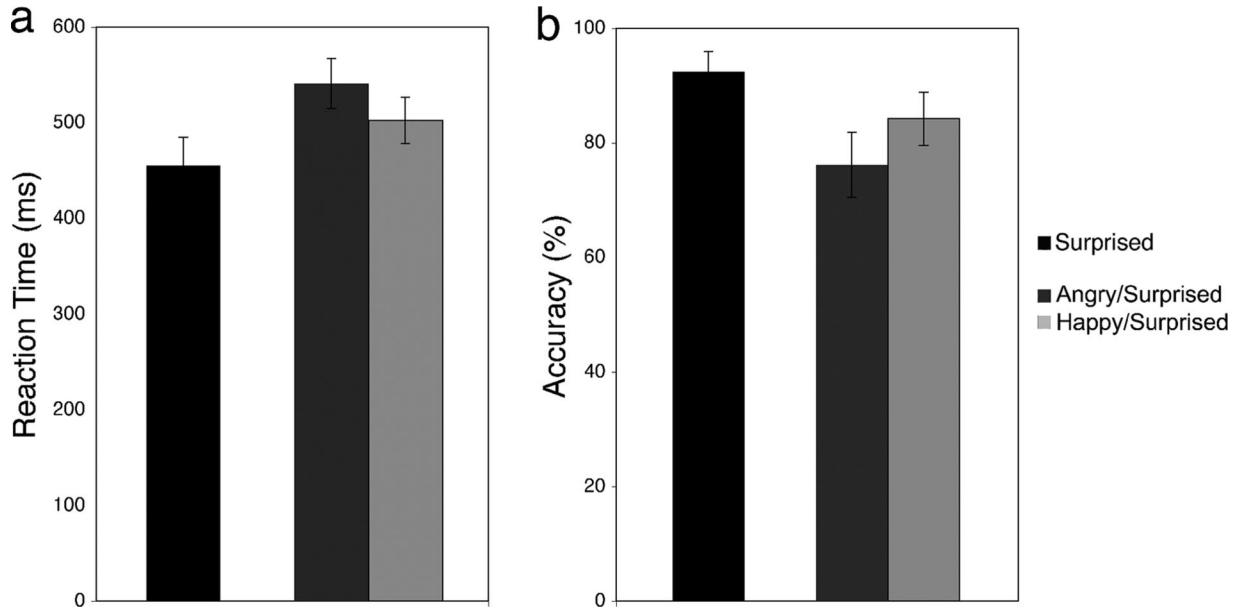


Figure 4. Initial negativity in surprised context. (a) Surprised faces, given their default negativity, represented a more negative context in which happy faces were detected faster than angry faces. (b) Participants were more accurate when detecting happy, as compared to angry faces, but that there was no difference in accuracy for angry and happy, as compared to surprised faces.

the default interpretation of surprised faces is negative. In order to test whether happy faces are detected faster than angry faces in any context, we ran a similar study where neutral faces serve as the context and found that there was no significant difference in reaction times for happy and angry faces (see Supplemental Materials, available online). This further supports the notion that surprised faces constitute an inherently negative context.

General Discussion

Here, we designed an emotional oddball task to provide evidence consistent with the notion that the default interpretation of surprise is negative (Experiments 2, 3). Specifically, participants were faster to detect surprised faces presented within a happy context, as compared to an angry context. Note that the emotional go/no go task uses a similar adaptation, but was designed to examine inhibition, rather than detection (Hare et al., 2005). Moreover, the present results examine the effect of implicit contexts (i.e., valence of standard faces of the oddball design in which surprised faces are infrequent targets) on the valence ratings of surprise (Experiment 1). We found that clearly valenced facial expressions can provide an implicit context that is sufficient to modulate valence interpretations of surprised faces in a congruent fashion. In other words, surprised faces were rated more positively when presented within a context of happy faces, and rated more negatively when presented within a context of angry faces. We discuss the implications of these findings within the scope of contrast effects and visual attention paradigms.

Uniqueness of Surprised Facial Expressions

Interestingly, the result that surprise is interpreted as congruent with the context (Experiment 1) is opposite that observed

previously to neutral faces (Russell & Fehr, 1987), where an implicit context of clearly valenced facial expressions caused an anchoring in the interpretation of neutral faces. Perhaps these differing outcomes are driven by inherent differences between surprised and neutral faces. Specifically, while neutral faces are intended to be devoid of emotional meaning, surprised faces can predict both positive and negative outcomes. This dual-valence representation likely renders surprise faces more context-dependent. Indeed, when we encounter surprised faces in our social world, the meaning of the expression is defined by the context (e.g., a birthday party, a traffic accident). Thus, it makes sense that the resolution of the predictive meaning of surprise faces should be congruent with contextual information.

Automaticity of Negative Ratings of Surprised Faces

In Experiment 1, we demonstrated that an implicit negative context had a greater influence on biasing ratings of surprise in a negative direction than the positive context did in a positive direction. This suggests that surprised faces are more inherently negative than positive. In Experiment 2, we found that participants were faster to detect surprised faces within a context of happy faces, as compared to a context of angry faces. We interpret this finding to mean that surprise is inherently more negative than positive and, thus, stood out as an “oddball” in the happy compared to angry context. Finally, Experiment 3 provided additional support for these findings by using surprised faces as the contextual stimuli and then measuring reaction times to angry or happy faces as the oddball stimuli. Reaction times to happy faces were faster than reaction times to angry faces—presumably because happy stood out as “oddballs” more in the context of surprised faces.

These findings are consistent with previous work showing that participants take longer to rate surprised faces as positive, as compared to ratings the same faces as negative (Neta et al., 2009). Moreover, this is consistent with previous data showing that a negative interpretation of surprised expressions is more elemental and more heavily represented in LSF information, while a positive interpretation of these expressions is more heavily dependent on HSF information (Neta & Whalen, 2010). Indeed, visual object recognition studies integrating functional MRI (fMRI) and magnetoencephalography (MEG) have shown that a coarse version of a visual stimulus, comprising mainly LSFs, is rapidly projected from early visual regions to the OFC, critically preceding activity in object processing regions in the inferior temporal cortex (IT; Bar et al., 2006; Kveraga et al., 2007). Thus, it has been suggested that LSF information within emotional stimuli may subserve an analogous “initial guess” process as part of a rapid threat detection or “early warning” system (de Gelder et al., 1999; LeDoux, 1996; Morris et al., 1999; Rafal et al., 1990; Sahraie et al., 2002). Further, previous findings demonstrated that the amygdala is particularly responsive to LSF versions of fearful faces (Vuilleumier et al., 2003), as well as to surprised faces that are interpreted negatively (Kim et al., 2003). These data support the idea that a brain circuitry that includes the amygdala may prioritize the processing of negativity over positivity. Additionally, previous research shows that a positivity bias requires an extra layer of processing in the form of additional prefrontal—amygdala interaction (Kim et al., 2003; Sharot et al., 2007).

Importantly, while Experiment 1 demonstrates that an implicit context can modulate an individual’s interpretation of surprise, Experiments 2 and 3 show that this context effect does not override the basic default negativity inherent to these expressions. As such, while we already demonstrated this default negativity in previous work (Neta & Whalen, 2010), we now show that this effect is (a) evident even when subjects are asked to perform a task that is unrelated to the specific valence evaluation of the face (i.e., detection), and (b) robust enough to be found when surprised expressions are placed within a context that helps to disambiguate their valence.

Relating the Present Data to Amygdala Responsivity

The amygdala has been shown to respond to unseen presentations of fearful faces (Jiang & He, 2006; Morris et al., 1998a; Morris et al., 1998b; Rauch et al., 2000; Whalen et al., 1998; Williams et al., 2004; Williams et al., 2006), and it has been shown to respond most strongly to LSFs of fearful faces (Vuilleumier et al., 2003). This is consistent with the assertion that the amygdala may prioritize the processing of negativity over the processing of positivity (LeDoux, 1996), especially in situations of predictive uncertainty (Kim et al., 2003; Whalen, 1998; 2007; Whalen et al., 2001). It should be noted that ample data in animal subjects and preliminary neuroimaging data in humans suggest that different portions of the amygdala (i.e., subnuclei) play different roles in the detection of negativity (Davis & Whalen, 2001; Kapp et al., 1992; LeDoux, 1996; LeDoux et al., 1990; Repa et al., 2001), positivity (Baxter & Murray, 2002; Gallagher et al., 1990; Paton et al., 2006; Radwanska et al., 2002), and/or the resolution of ambiguity (Davis et al., 2010; Kim et al., 2003; Whalen et al., 2001).

With specific relevance to surprised faces, activity in the dorsal amygdaloid region increased across all subjects regardless of whether they interpreted surprised faces as positive or negative (i.e., an arousal-based response). Conversely, responses within the ventral amygdala were consistent with the ascribed valence of the faces, and critically were higher to negative interpretations compared to positive interpretations (Kim et al., 2003). These data suggest that one portion of the human amygdala may show more automated responses where processing of negativity is prioritized (e.g., ventral amygdala where BLA is located), while another portion of the amygdala is equally sensitive to predictors of negative and positive events, particularly when the outcome predicted by these events (i.e., surprised expression) remains ambiguous (e.g., dorsal amygdala/substantia innominata where central nucleus is located, see Whalen et al., 2009 for discussion).

Conclusion

Here we offer three experiments demonstrating that (a) implicit contextual information consisting of other facial expressions modulates valence assessments of surprised faces, (b) this modulation causes surprised faces to be interpreted as congruent with the valence of the contextual expressions, and (c) reaction times in an oddball paradigm are consistent with the notion that the default valence interpretation of surprise is negative. The implication is that interpreting the ambiguous facial expression of surprise in a more positive light will require a bit more effort, since the default interpretation will need to be regulated and overwritten. These tasks could be utilized for the study of anxiety disorders and major depressive disorder, since these subjects tend to show a negativity bias when assessing facial expressions of emotion (Mogg & Bradley, 2005; Norris et al., 2009; Williams et al., 2007).

References

- Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmid, A. M., Dale, A. M., . . . Halgren, E. (2006). Top-down facilitation of visual recognition. *Proceedings of the National Academy of Sciences*, 103, 449–454.
- Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: An investigation of adults with asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders*, 34, 163–175.
- Baxter, M., & Murray, E. (2002). The amygdala and reward. *Nature Reviews Neuroscience*, 3, 563–573.
- Beck, A. T., Ward, C., & Mendelson, M. (1961). Beck Depression Inventory (BDI). *Archives of General Psychiatry*, 4, 561–571.
- Bouton, M. E. (1994). Context, ambiguity, and classical conditioning. *Current Directions in Psychological Science*, 3, 49–53.
- Campanella, S., Gaspard, C., Debatisse, D., Bruyer, R., Crommelinck, M., & Guerit, J.-M. (2002). Discrimination of emotional facial expressions in a visual oddball task: An ERP study. *Biological Psychology*, 59, 171–186.
- Canli, T., Sivers, H., Whitfield, S. L., Gotlib, I. H., & Gabrieli, J. D. E. (2002). Amygdala response to happy faces as a function of extraversion. *Science*, 296, 2191.
- Costa, P. T., & McCrae, R. R. (1991). Neo Five-Factor Inventory (NEO-FFI) Professional Manual. Psychological Assessment Resources: Odessa, FL.
- Davis, F. C., Johnstone, T., Mazzulla, E. C., Oler, J. A., Whalen, P. J. (2010). Regional response differences across the human amygdaloid complex during social conditioning. *Cerebral Cortex*, 20, 612–621.

- Davis, M., & Whalen, P. J. (2001). The amygdala: Vigilance and emotion. *Molecular Psychiatry*, 6, 13–34.
- de Gelder, B., Vroomen, J., Pourtois, G., & Weiskrantz, L. (1999). Non-conscious recognition of affect in the absence of striate cortex. *NeuroReport*, 10, 3759–3763.
- Ekman, P., & Friesen, W. V. (1976). Pictures of Facial Affect [slides]. Palo Alto, CA: Consulting Psychologists Press.
- Gallagher, M., Graham, P. W., & Holland, P. C. (1990). The amygdala central nucleus and appetitive Pavlovian conditioning: Lesions impair one class of conditioned behavior. *Journal of Neuroscience*, 10, 1906–1911.
- Hare, T. A., Tottenham, N., Davidson, M. C., Glover, G. H., & Casey, B. J. (2005). Contributions of amygdala and striatal activity in emotion regulation. *Biological Psychiatry*, 57, 624–632.
- Hughes, H. C., Nozawa, G., & Kitterle, F. (1996). Global precedence, spatial frequency channels, and the statistics of natural images. *Journal of Cognitive Neuroscience*, 8, 197–230.
- Jiang, Y., & He, S. (2006). Cortical responses to invisible faces: Dissociating subsystems for facial-information processing. *Current Biology*, 16, 2023–2029.
- Kapp, B. S., Whalen, P. J., Supple, W. F., & Pascoe, J. P. (1992). Amygdaloid contributions to conditioned arousal and sensory information processing. In J. P. Aggleton (Ed.), *The Amygdala: Neurobiological Aspects of Emotion, Memory, and Mental Dysfunction* (pp. 229–254). New York: Wiley-Liss.
- Kim, H., Somerville, L. H., Johnstone, T., Alexander, A., & Whalen, P. J. (2003). Inverse amygdala and medial prefrontal cortex responses to surprised faces. *Neuroreport*, 14, 2317–2322.
- Kim, H., Somerville, L. H., Johnstone, T., Polis, S., Alexander, A. L., Shin, L. M., & Whalen, P. J. (2004). Contextual modulation of amygdala responsivity to surprised faces. *Journal of Cognitive Neuroscience*, 16, 1730–1745.
- Kveraga, K., Boshyan, J., & Bar, M. (2007). Magnocellular projections as the trigger of top-down facilitation in recognition. *Journal of Neuroscience*, 27, 13232–13240.
- LeDoux, J. E. (1996). The Emotional Brain. Simon & Schuster, New York.
- LeDoux, J. E., Cicchetti, P., Xagoraris, A., & Romanski, L. M. (1990). The lateral amygdaloid nucleus: Sensory interface of the amygdala in fear conditioning. *Journal of Neuroscience*, 10, 1062–1069.
- Mogg, K., & Bradley, B. P. (2005). Attentional bias in generalized anxiety disorder versus depressive disorder. *Cognitive Therapy and Research*, 29, 29–45.
- Morris, J. S., Friston, K. J., Buchel, C. D., Young, A. W., Calder, A. J., & Dolan, R. J. (1998b). A neuromodulatory role for the human amygdala in processing emotional facial expressions. *Brain*, 121, 47–57.
- Morris, J. S., Ohman, A., & Dolan, R. J. (1998a). Conscious and unconscious emotional learning in the human amygdala. *Nature*, 393, 467–470.
- Morris, J. S., Ohman, A., & Dolan, R. J. (1999). A subcortical pathway to the right amygdala mediating “unseen” fear. *Proceedings of the National Academy of Sciences*, 96, 1680–1685.
- Neta, M., Norris, C. J., & Whalen, P. J. (2009). Corrugator muscle responses are associated with individual differences in positivity-negativity bias. *Emotion*, 9, 640–648.
- Neta, M., & Whalen, P. J. (2010). The primacy of negative interpretations when resolving the valence of ambiguous facial expressions. *Psychological Science*, 21, 901–907.
- Norris, C. J., Gollan, J., Berntson, G. G., & Cacioppo, J. T. (2009). *The Current Status of Research on the Structure of Affective Space*. Invited manuscript submitted to Biological Psychology: Special Issue on Emotion.
- Paton, J. J., Belova, M. A., Morrison, S. E., & Salzman, C. D. (2006). The primate amygdala represents the positive and negative value of visual stimuli during learning. *Nature*, 439, 865–870.
- Radwanska, K., Nikolaev, E., Knapska, E., & Kaczmarek, L. (2002). Differential response of two subdivisions of lateral amygdala to aversive conditioning as revealed by c-Fos and P-ERK mapping. *Neuroreport*, 13, 2241–2246.
- Rafal, R. Smith, J., Krantz, J., Cohen, A., & Brennan, C. (1990). Extra-geniculate vision in hemianopic humans: Saccade inhibition by signals in the blind field. *Science*, 250, 118–121.
- Rauch, S. L., Whalen, P. J., Shin, L. M., McInerney, S. C., Orr, S., Lasklo, N., & Pitman, R. (2000). Exaggerated amygdala response to masked facial expressions in posttraumatic stress disorder. *Biological Psychiatry*, 47, 769–776.
- Repa, J. C., Muller, J., Apergis, J., Desrochers, T. M., Zhou, Y., & LeDoux, J. E. (2001). Two different lateral amygdala cell populations contribute to the initiation and storage of memory. *Nature Neuroscience*, 4, 724–731.
- Russell, J. A., & Fehr, B. (1987). Relativity in the perception of emotion in facial expressions. *Journal of Experimental Psychology: General*, 116, 223–237.
- Sahraie, A., Weiskrantz, L., Trevethan, C. T., Cruce, R., & Murray, A. D. (2002). Psychophysical and pupillometric study of spatial channels of visual processing in blindsight. *Experimental Brain Research*, 143, 249–259.
- Sharot, T., Riccardi, A. M., Raio, C. M., & Phelps, E. A. (2007). Neural mechanisms mediating optimism bias. *Nature*, 450, 102–105.
- Somerville, L. H., Kim, H., Johnstone, T., Alexander, A. L., & Whalen, P. J. (2004). Human amygdala responses during presentation of happy and neutral faces: Correlations with state anxiety. *Biological Psychiatry*, 55, 897–903.
- Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1988). STAI - Manual for the State Trait Anxiety Inventory. Consulting Psychologists Press: Palo Alto, CA.
- Tottenham, N., Tanaka, J., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., . . . Nelson, C. A. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research*, 168, 242–249.
- Tsuchiya, N., Moradi, F., Felsen, C., Yamazaki, M., Adolphs, R. (2009). Intact rapid detection of fearful faces in the absence of the amygdala. *Nature Neuroscience*, 12, 1224–1225.
- Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2003). Distinct spatial frequency sensitivities for processing faces and emotional expressions. *Nature Neuroscience*, 6, 624–631.
- Whalen, P. J. (1998). Fear, vigilance, and ambiguity: Initial neuroimaging studies of the human amygdala. *Current Directions in Psychological Science*, 7, 177–188.
- Whalen, P. J. (2007). The uncertainty of it all. *Trends in Cognitive Sciences*, 11, 499–500.
- Whalen, P. J., Davis, F. C., Oler, J. A., Kim, H., Kim, M. J., & Neta, M. (2009). Human amygdala responses to facial expressions of emotion. In E. A. Phelps & P. J. Whalen (Eds.), *The Human Amygdala*. New York: Guilford Press.
- Whalen, P. J., Rauch, S. L., Etcoff, N. L., McInerney, S. C., Lee, M., & Jenike, M. A. (1998). Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *Journal of Neuroscience*, 18, 411–418.
- Whalen, P. J., Shin, L. M., McInerney, S. C., & Fischer, H. (2001). A functional MRI study of human amygdala responses to facial expressions of fear versus anger. *Emotion*, 1, 70–83.
- Williams, L. M., Kemp, A. H., Felmingham, K., Liddell, B. J., Palmer, D. M., & Bryant, R. A. (2007). Neural biases to covert and overt signals of fear: Dissociation by trait anxiety and depression. *Journal of Cognitive Neuroscience*, 19, 1595–1608.
- Williams, L. M., Liddell, B. J., Kemp, A. H., Bryant, R. A., Meares, R. A., Peduto, A. S., & Gordon, E. (2006). Amygdala-prefrontal

- dissociation of subliminal and supraliminal fear. *Human Brain Mapping*, 27, 652–661.
- Williams, M. A., Morris, A. P., McGlone, F., Abbott, D. F., & Mattingley, J. B. (2004). Amygdala responses to fearful and happy facial expressions under conditions of binocular rivalry. *Journal of Neuroscience*, 24, 2898–2904.
- Wyer, N. A., Sherman, J. W., & Stroessner, S. J. (2000). The roles of motivation and ability in controlling the consequences of stereotype suppression. *Personality & Social Psychology Bulletin*, 26, 13–25.

Received September 8, 2010

Revision received January 7, 2011

Accepted January 7, 2011 ■