Social Cognitive and Affective Neuroscience Advance Access published February 29, 2016

Social Cognitive and Affective Neuroscience, 2016, 1–8

doi: 10.1093/scan/nsw003
Original article

Interpreting ambiguous social cues in unpredictable contexts

F. Caroline Davis,[1] Maital Neta,[2] M. Justin Kim,[3] Joseph M. Moran,[1,4] and Paul J. Whalen[3]

[1]Cognitive Science, US Army Natick Soldier Research Development and Engineering Center (NSRDEC), Natick, MA, [2]Center for Brain, Biology, and Behavior, University of Nebraska-Lincoln, Lincoln, NE, [3]Department of Psychological and Brain Sciences, Dartmouth College, Hanover, NH, and [4]Department of Psychology, Harvard University, Cambridge, MA, USA

Correspondence should be addressed to F. Caroline Davis, Natick Soldier Research Development and Engineering Center(NSRDEC), 10 Kansas St, Natick, MA 01760, USA. E-mail: f.c.davis4.civ@mail.mil.

Abstract

Unpredictable environments can be anxiety-provoking and elicit exaggerated emotional responses to aversive stimuli. Even neutral stimuli, when presented in an unpredictable fashion, prime anxiety-like behavior and elicit heightened amygdala activity. The amygdala plays a key role in initiating responses to biologically relevant information, such as facial expressions of emotion. While some expressions clearly signal negative (anger) or positive (happy) events, other expressions (e.g. surprise) are more ambiguous in that they can predict either valence, depending on the context. Here, we sought to determine whether unpredictable presentations of ambiguous facial expressions would bias participants to interpret them more negatively. We used functional magnetic resonance imaging and facial electromyography (EMG) to characterize responses to predictable *vs* unpredictable presentations of surprised faces. We observed moderate but sustained increases in amygdala reactivity to predictable presentations of surprised faces, and relatively increased amygdala responses to unpredictable faces that then habituated, similar to previously observed responses to clearly negative (e.g. fearful) faces. We also observed decreased corrugator EMG responses to predictable surprised face presentations, similar to happy faces, and increased responses to unpredictable surprised face presentations, similar to angry faces. Taken together, these data suggest that unpredictability biases people to interpret ambiguous social cues negatively.

Key words: amygdala; emotion; facial expression; Unpredictability

Introduction

Unpredictability is inherently aversive. Noxious stimuli become even more aversive when embedded within an unpredictable context (Weiss, 1970; Mineka and Kihlstrom, 1978; Grillon *et al.*, 2004, 2006; Vansteenwegen *et al.*, 2008). Even emotionally neutral stimuli (e.g. tones), when presented in an unpredictable fashion, bias organisms toward hypervigilance for threat and potentiate anxiety-like behavior (Herry *et al.*, 2007). Unpredictable environments have been shown to influence sympathetic nervous system reactivity (Davies and Craske, 2015), and to facilitate cognitive processes such as early sensory processing (Shackman *et al.*, 2011) and the ability to detect stimuli that predict threat (Robinson *et al.*, 2011).

Brain regions implicated in fear learning and threat detection, such as the amygdala, are also sensitive to unpredictability (Herry *et al.*, 2007; Whalen, 2007; Alvarez *et al.*, 2011). Associative learning studies show that the amygdala is involved in orienting toward a surprising (unpredicted) stimulus and then initiating behaviors and neural processes that facilitate learning about the environment (Kapp *et al.*, 1992; Whalen, 1998; Holland

Received: 30 June 2015; **Revised:** 8 October 2015; **Accepted:** 4 January 2016

and Gallagher, 1999). The amygdala is also highly responsive to facial expressions of emotion (Costafreda *et al.*, 2008; Sergerie *et al.*, 2008), presumably because they are social cues that predict biologically relevant events such as threat or reward (Whalen, 1998).

While the signals conveyed by some facial expressions are relatively clear (e.g. an angry facial expression predicts aggressive behavior), others are more ambiguous. For example, fearful facial expressions communicate threat, but do not provide much information about what or where that threat is (Whalen, 1998). Surprised facial expressions are even more ambiguous in that they signal an important event, but provide little information about whether that event is threatening or rewarding. In order to properly respond to a surprising event or stimulus, an individual must use cues in the environment to resolve this ambiguity. Indeed, both valence interpretations and amygdala responses to surprised faces are highly sensitive to contextual information (Kim *et al.*, 2004; Neta *et al.*, 2011). However, when presented in the absence of a clarifying context, participants will default to their 'spontaneous' interpretation of ambiguous facial expressions, which varies widely across individuals (Kim *et al.*, 2003; Neta *et al.*, 2009).

During every day social interactions, contextual information can take many forms including verbal cues, body language or environmental factors. Here, we sought to determine whether unpredictable presentations of surprised faces would influence the way that individuals interpret those faces. We hypothesized that if (i) unpredictability can serve as a contextual cue that biases organisms toward hypervigilance for threat and (ii) valence interpretations of ambiguous social cues are context dependent, then presenting ambiguous (surprised) facial expressions in a temporally unpredictable context will bias people to interpret them negatively. In Experiment 1, we used functional magnetic resonance imaging (fMRI) to characterize neural responses to predictable *vs* unpredictable presentations of surprised faces. Specifically, we hypothesized that temporal unpredictability would elicit heightened amygdala responses to the faces; consistent with patterns of amygdala reactivity to clearly negative facial expressions (e.g. fear). Since the amygdala is not uniquely responsive to negative stimuli (Holland and Gallagher, 1999; Paton *et al.*, 2006; Costafreda *et al.*, 2008; Sergerie *et al.*, 2008), we conducted a second experiment using facial electromyography (EMG), which is considered an objective measure of valence interpretations of emotional stimuli (Cacioppo *et al.*, 1986), to further test the hypothesis that unpredictability biases individuals to interpret ambiguous cues negatively. Specifically, we examined corrugator responses to predictable and unpredictable presentations of surprised, angry and happy faces. We predicted that corrugator activity would be enhanced during unpredictable presentations of surprised faces (similar to angry faces) and decreased in response to predictable surprised faces (similar to happy faces).

Experiment 1: Materials and methods

Participants

Twenty-three healthy adult participants were recruited for this experiment. One participant was excluded for excessive movement (i.e.>2 mm), leaving a total of 22 participants (12 males, ages 18–40) in the fMRI analyses. All participants were right-handed as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971) and free from psychiatric, neurological and medical illness. Psychiatric history was assessed using an abbreviated version of the Non-Patient edition of the Structured Clinical Interview for DSM Disorders (First *et al.*, 1995), which assessed for current and past history of major depressive disorder, dysthymia, hypomania, bipolar disorder, specific phobia, social anxiety disorder, generalized anxiety disorder and obsessive compulsive disorder. Neurological and medical histories were assessed through self-report. This investigation was conducted in accordance with the guidelines of the Human Subjects Committee of the University of Wisconsin-Madison.

Stimuli

Visual stimuli were presented using E-Prime software (Psychology Software Tools, Pittsburgh, PA) through a fiber-optic goggle system (AVOTEC, Stuart, FL) mounted on a quadrature head coil. Face stimuli (Tottenham *et al.*, 2009) were eight surprised faces (four male). Faces were aligned vertically with the middle of the pupils and horizontally with the nose in the center, and matched on contrast and brightness. Faces were presented in the center of the computer screen on a black background, and fixation crosshairs were white on a black background.

Procedure

Because amygdala responses to emotional stimuli tend to be more robust during passive compared to task-based paradigms (Costafreda *et al.*, 2008), we employed a passive viewing paradigm where participants were told that they would see a series of pictures and that they should watch the faces carefully and focus on the fixation crosshair between trials. All participants viewed two runs of surprised face presentations, each consisting of 16-second blocks of fixation alternating with 16-second blocks of faces (Figure 1). Within each run, participants viewed three blocks of Predictable (P) face presentations (i.e. 16 faces presented for 500 ms each, alternating with 500 ms of fixation), and three blocks of Unpredictable (U) face presentations (i.e. 16 faces presented for 500 ms each, alternating with randomly jittered fixation lasting either 300, 500 or 700 ms). The order of stimulus presentation blocks (i.e. PUPUPU *vs* UPUPUP) was counterbalanced across participants. This task was run as a separate study in conjunction with, and always following, a different study that involved presentation of neutral faces only.

Image acquisition

Images were acquired on a 3.0 Tesla MRI scanner (General Electric SIGNA; Waukesha, WI) with high speed imaging gradients and a quadrature head coil. Anatomical images were whole brain high resolution T1-weighted scans (3D Inversion Recovery; 256 × 256 in-plane resolution; 240 mm FOV; 124 × 1.1 mm axial slices). Functional scans consisted of an EPI sequence with a 2000 ms repetition time (TR), 33 ms echo time (TE), and a 60° flip angle with 18 contiguous 3 mm thick interleaved axial slices (0.5 mm interslice gap; 64 × 64 in-plane resolution, 180 mm FOV) oriented parallel to the Anterior Commissure—Posterior Commissure line.

Data analysis

All fMRI data analyses were performed using SPM8 (Wellcome Department of Cognitive Neurology, London, UK). Raw functional data were realigned to the mean image, segmented, coregistered and spatially normalized into standard (MNI) space and resampled to 2 × 2 × 2 mm voxels. Normalized functional data

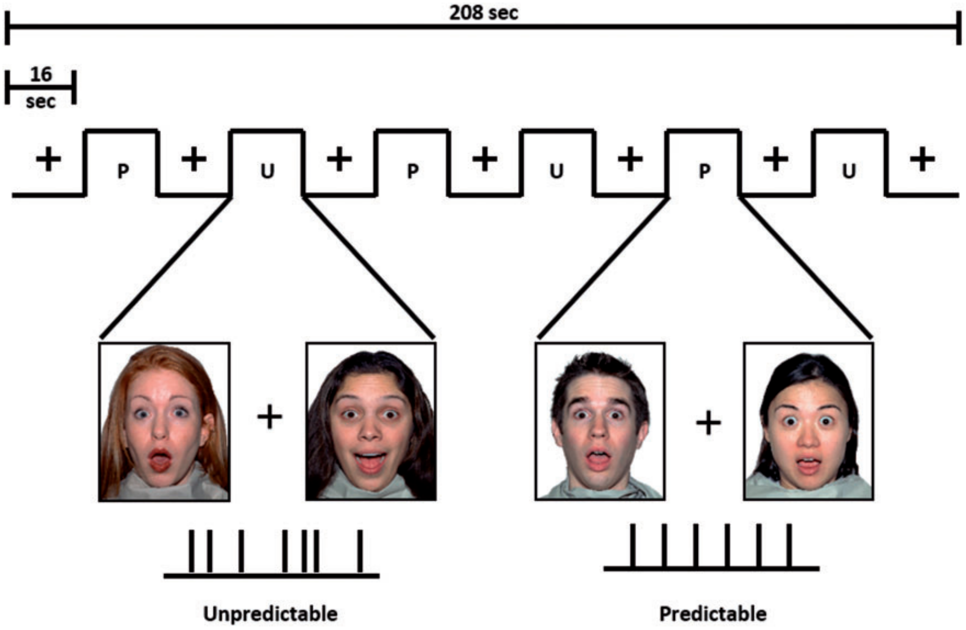

Fig. 1. Experimental Design. Faces were presented in 16 second blocks, alternating with 16 second blocks of fixation. Blocks alternated between 'Unpredictable' and 'Predictable' presentations of surprised faces.

were then smoothed using a Gaussian kernel of 6 mm full width at half maximum.

Since amygdala responses tend to habituate over time (Buchel *et al.*, 1998, 1999; Labar *et al.*, 1998; Morris *et al.*, 2001; Phelps *et al.*, 2001; Cheng *et al.*, 2007; Plichta *et al.*, 2014), we assessed variation in habituation rates as a function of predictability by comparing Early and Late responses. A general linear model employing the canonical hemodynamic response function in SPM8 was fitted to signal subject data with four regressors of interest: Early Predictable (Run 1), Late Predictable (Run 2), Early Unpredictable (Run 1) and Late Unpredictable (Run 2). Linear contrast images were generated and then entered into a second-level random effects model to determine mean BOLD responses for each condition using one-sample *t*-tests with a voxel level statistical threshold of $P < 0.05$, familywise error (FWE) corrected for multiple comparisons across the whole brain.

To investigate the amygdala's response to Predictable and Unpredictable face presentations in an unbiased manner, we then restricted our search volume to functional clusters that exceeded our correction criteria in the All Faces *vs* Baseline contrast within the left and right amygdala, defined using the AAL atlas (Tzourio-Mazoyer *et al.,* 2002) in the WFU PickAtlas toolbox (Maldjian *et al.*, 2003, 2004). We identified the peak voxels within these clusters, and then extracted average signal estimates from spherical regions of interest (ROIs) (radius = 3 mm) surrounding these peak voxels. These values were then submitted to a 2 (Time: Early, Late) × 2 (Predictability: Predictable, Unpredictable) repeated measures ANOVA using IBM SPSS Statistics 17.0. Directionality of significant interactions was evaluated using paired samples *t*-tests. All coordinates are reported in MNI space.

Experiment 1: Results

A 2 (Time: Early, Late) × 2 (Predictability: Predictable, Unpredictable) within subjects ANOVA performed on signal estimates extracted from the left amygdala (peak voxel: $x = -22$, $y = -4$, $z = -18$) and right amygdala (peak voxel: $x = 20$, $y = -4$, $z = -18$) ROIs yielded no main effect of Predictability (left: F(1,21) = 0.182, $P = 0.674$; right: F(1,21) = 0.391, $P = 0.538$), a marginally significant main effect of Time in the left amygdala (F(1,21) = 3.857, $P = 0.063$) and a significant main effect of Time in the right amygdala(F(1,21) = 7.837, $P = 0.011$). We also observed a significant interaction between Time and Predictability in the left amygdala (F(1,21) = 10.163, $P = 0.004$), and a marginally significant interaction in the right amygdala (F(1,21) = 3.814, $P = 0.064$). To further explore the significant Time x Predictability interaction in the left amygdala, we computed paired samples *t*-tests on the means relevant to our hypotheses. We observed no significant difference between Predictable Early and Predictable Late (t(21) = .227, $P = 0.823$). In contrast, there was a significant difference between Unpredictable Early and Unpredictable Late, such that the amygdala showed greater activity in Early compared to Late trials in the Unpredictable condition (t(21) = 3.277, $P = 0.004$). In addition, Early amygdala responses to Unpredictable faces were greater than those to Predictable faces (t(21) = 2.470, $P = 0.022$) and Late amygdala responses to Unpredictable faces were less than those to Predictable faces (t(21) = −2.227, $P = 0.037$) (Figure 2). FDR correction for these four *t*-tests yielded a corrected threshold of $P < 0.0375$, which corresponds to $q < 0.05$, and all three significant comparisons survive this correction (Benjamini and Hochberg, 1995).

Experiment 1: Discussion

Amygdala responses to predictable surprised faces were of relatively moderate amplitude and sustained over time, whereas amygdala responses to unpredictable surprised faces were enhanced compared with predictable faces, but this response habituated over time. Interestingly, these data are similar to other data from our lab, where amygdala responses to blocks of surprised faces (presented with predictable timing) were moderate but sustained over time, whereas responses to fearful facial

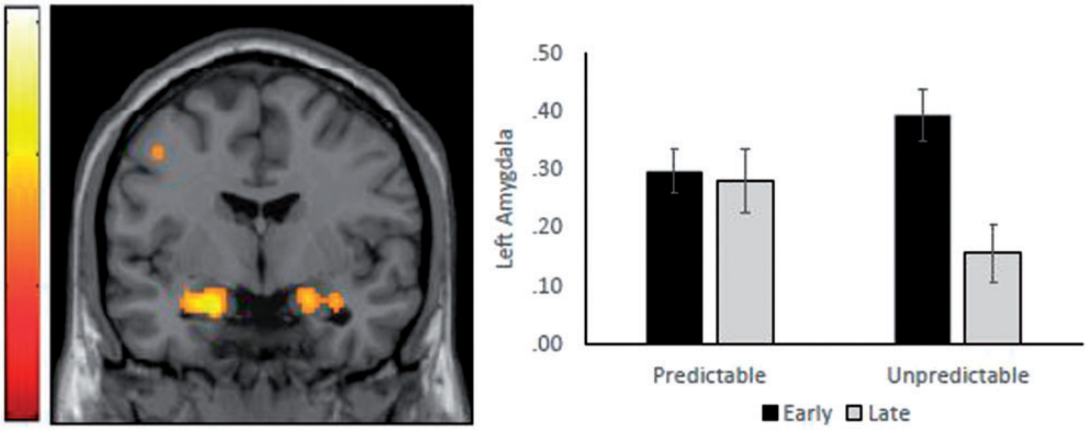

Fig. 2. Amygdala reactivity to all faces compared to baseline. (A) Right amygdala peak: x = 20, y = −4, z = −18. Left amygdala peak: x = −22, y = −4, z = −18. Statistical map thresholded at $P < 0.05$, FWE. (B) Signal estimates extracted from a spherical ROI with a 3 mm radius drawn around the left amygdala peak coordinate (Note: Right amygdala signal estimates not shown, but pattern mirrors that found in the left amygdala). Error bars represent standard error of the mean.

expressions (also presented with predictable timing) were initially enhanced compared to surprised face blocks but habituated over time (Kim *et al.*, 2003; see Whalen *et al.*, 2009, see p. 275, Plate 12.5). In other words, while the observed amygdala responses to predictable surprised expressions in the present study were similar to previous data exploring amygdala response to surprised faces (with no additional contextual information), responses to unpredictable surprised expressions were more similar to amygdala responses to clearly negative fearful facial expressions. In light of these data, we suggest that the patterns of activity observed here were due to the fact that the unpredictable context was sufficient to bias participants to interpret surprised faces negatively.

While the human amygdala is highly responsive to clearly negative stimuli, it also responds to positive stimuli (Holland and Gallagher, 1999; Paton *et al.*, 2006; Costafreda *et al.*, 2008; Sergerie *et al.*, 2008). Therefore, an amygdala response in and of itself is not evidence that the presented stimulus was necessarily interpreted in a negative light. Thus, in Experiment, 2 we used an objective measure of implicit valence interpretations of emotional stimuli (i.e. changes in corrugator supercilii activity measured using facial EMG; Cacioppo *et al.*, 1986) to test our hypothesis that participants were interpreting unpredictable surprised faces negatively.

Experiment 2: Materials and methods

Participants

Fifty-one healthy adult participants were recruited for this experiment. Five participants were excluded because of technical problems during data collection, leaving 46 participants (15 males, ages 18–40) in this analysis. All participants were free from psychiatric, neurological and medical illness. Psychiatric history was assessed using an abbreviated version of the Non-Patient edition of the Structured Clinical Interview for DSM Disorders (First *et al.*, 1995), which assessed participants for current and past history of major depressive disorder, dysthymia, hypomania, bipolar disorder, specific phobia, social anxiety disorder, generalized anxiety disorder and obsessive compulsive disorder. Neurological and medical histories were assessed through self-report. This investigation was conducted in accordance with the guidelines of the Human Subjects Committee of Dartmouth College.

Stimuli

Visual stimuli were presented using E-Prime software (Psychology Software Tools, Pittsburgh, PA). Face stimuli (Tottenham *et al.*, 2009) consisted of eight surprised (four male) faces, eight happy (four male) and eight angry (four male) faces. Faces were aligned vertically with the middle of the pupils and horizontally with the nose in the center, and matched on contrast and brightness. Faces were presented in the center of the computer screen on a black background, and fixation crosshairs were white on a black background.

Procedure

Consistent with the methods of Experiment 1, we used a passive viewing paradigm, where participants were seated in a dimly lit room and told that they would view a series of pictures. They were instructed to watch the faces carefully and focus on the fixation crosshair between trials. All participants viewed two runs of surprised face presentations, as in Experiment 1 (Figure 1). Next, participants viewed two runs, each consisting of alternating blocks of Happy (H) and Angry (A) faces. This order was fixed (i.e. surprise runs followed by Happy/Angry runs) so that the clearly valenced happy and angry faces did not influence participants' implicit valence interpretations of the surprised faces. Indeed, we have previously shown that valence interpretations of surprise can be modulated by a context of these clearly valenced expressions (angry, happy; Neta *et al.*, 2011). Happy and Angry face presentations were all 'Predictable', with a fixation cross lasting 500 ms separating the faces. The presentation order of Happy and Angry blocks (i.e. HAHAHA *vs* AHAHAH) was counterbalanced across participants.

During each testing session, a female experimenter attached electrodes for recording facial EMG according to standard methods (for details, see Fowles *et al.*, 1981; Fridlund and Cacioppo, 1986). To allow participants to relax and acclimate to the testing environment before beginning the experimental trials, participants viewed 27 images of neutral scenes (e.g. landscapes), during which they were asked to press a button corresponding to whether the scene was indoors or outdoors.

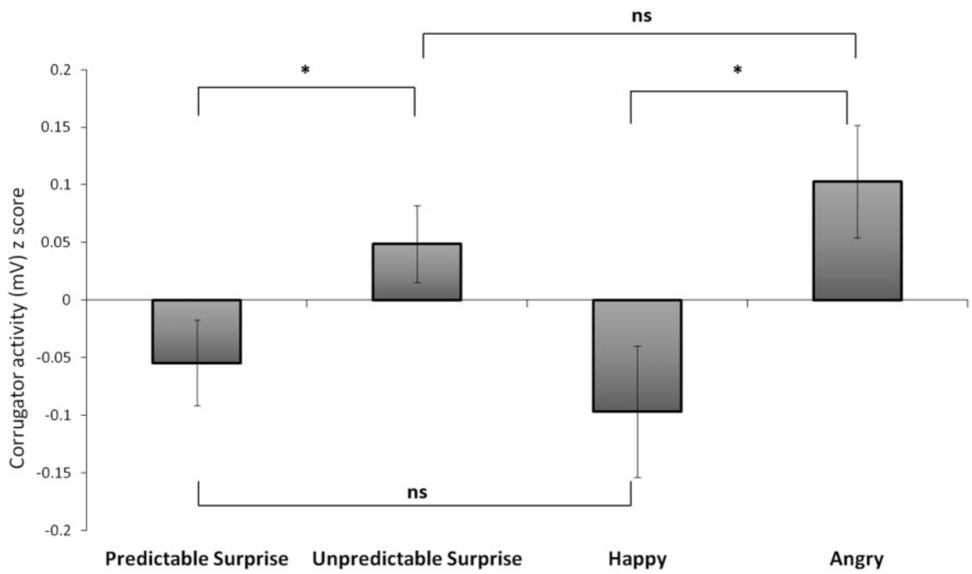

Fig. 3. Corrugator responses to Unpredictable surprised faces were potentiated, similar to clearly negative, or Angry, expressions. In contrast, corrugator responses to Predictable surprised faces were attenuated, similar to clearly positive, or Happy, expressions. Error bars represent standard error of the mean. $^{*}P < 0.05$.

Physiological parameters and data reduction

EMG data were sampled and recorded at 1000 Hz. Data were converted and amplified with an eight-channel amplifier (PowerLab 8/30; ADInstruments, New South Wales, Australia) and displayed, stored, reduced and analyzed with the Chart 5.4.2 software package (ADInstruments, 2002).

Facial EMG was measured using 4-mm standard Ag/Ag-Cl electrodes, filled with Signa Gel electrode paste and attached bipolarly over the corrugator supercilii muscle region (Fridlund and Cacioppo, 1986). The skin was first cleaned with alcohol and rubbed with an abrasive gel. Electrode placement was counterbalanced such that half of the participants had electrodes placed on the right side of the face, and the other half on the left side of the face. To conceal the recording of facial muscle activity, we used a cover story, telling participants that we were measuring sweat gland activity.

Offline, EMG data were submitted to a DC Restore to center the signal at a zero point, a 60-Hz notch filter, a 30-Hz high-pass filter to reduce movement and blink-related artifact, then full rectified. To correct for the positive skew found in our EMG data (as has been previously found to be inherent to EMG data), all data were then subjected to a square-root transformation (Fridlund and Cacioppo, 1986). Corrugator activity was scored in 1000 ms increments and then averaged across the entire block, which lasted 16 000 ms after the first stimulus onset, and a baseline of 3000 ms, ending 500 ms before each block of faces began, resulting in an overall EMG change score. We used a longer baseline here than in previous work (Neta *et al.*, 2009) to be consistent with research that used facial EMG to assess state changes induced by emotional stimuli using a block design paradigm, similar to the one employed here (Smith *et al.*, 2005)

Because there are inherent individual differences in the range of corrugator reactivity to emotional stimuli (Cacioppo *et al.*, 1992; Lang *et al.*, 1993; Norris *et al.*, 2007), and our goal was not to evaluate individual differences but rather to compare EMG responses to the different experimental conditions, we standardized (i.e. z scored) responses for each participant (Bush *et al.*, 1993; Bensafi *et al.*, 2003). Previous work has analyzed data in this manner (Pattyn *et al.*, 2008; Neta *et al.*, 2009), as such interindividual variability is a common issue in psychophysiology (Bush *et al.*, 1993). By standardizing our data, we were able to calculate relative EMG responses to Predictable and Unpredictable Surprised faces, and Angry and Happy faces, for each individual, and then we can compare this relative activity across participants. These *z* scores were calculated based on the mean and standard deviation of activity across all trials for each participant, regardless of condition. We then averaged EMG activity across all trials within each condition, and computed a one way repeated measures ANOVA to evaluate any differences between the conditions (Predictable Surprise, Unpredictable Surprise, Happy, and Angry). Specific differences between conditions were evaluated using paired sample *t*-tests.

Experiment 2: Results

A one-way repeated measures ANOVA (Predictable Surprise, Unpredictable Surprise, Happy and Angry) revealed a significant difference between conditions ($F(3,43) = 3.490$, $P = 0.024$). In accordance with our predictions, paired sample *t*-tests revealed a significant difference in corrugator activity such that responses to Unpredictable presentations of Surprised faces were greater than those to Predictable faces ($t(45) = 2.097$, $P = 0.042$). Moreover, increases in corrugator activity during Unpredictable presentations of Surprised faces were not significantly different from increases in activity during Angry face blocks ($t(45) = -0.779$, $P = 0.440$), and decreases in activity during Predictable presentations of Surprised faces were not significantly different from decreases in activity during Happy face blocks ($t(45) = 0.502$, $P = 0.618$) (Figure 3).

Experiment 2: Discussion

In Experiment 2, we found that presenting surprised faces in a temporally unpredictable context biased participants to interpret the faces more negatively than those presented in a predictable fashion. Specifically, corrugator responses to unpredictable surprised faces were potentiated, similar to clearly negative (angry) expressions. Interestingly, corrugator responses to predictable surprised faces were attenuated, similar to clearly positive (happy) expressions. Perhaps in this experimental context, the contrast with unpredictable surprised faces biased participants to interpret the predictable presentations of surprised faces positively.

General Discussion

In this study, we manipulated the presentation rate of ambiguous social cues (i.e. surprised facial expressions) and found that unpredictable faces were interpreted more negatively than faces presented with predictable timing. These data build upon previous research by Herry et al. (2007), who found that listening to unpredictable tones primed anxiety-like behavior while performing unrelated tasks. Here, we extend those results using socially relevant stimuli to show that unpredictability can also influence the way that participants interpret the stimuli themselves, particularly when those stimuli do not have clear intrinsic valence (i.e. they are ambiguous).

In Experiment 1, we found that amygdala responses to unpredictable surprised faces were greater than responses to predictable surprised faces but habituated over time (Figure 2). This pattern of reactivity is consistent with previous work (Kim, 2003; Whalen *et al.*, 2009), where participants viewed blocks of surprised facial expressions and separate blocks of fearful facial expressions, both presented with predictable timing. In that experimental context, the more clearly negative (i.e. fearful) faces elicited initial amygdala responses of high magnitude that habituated rapidly, whereas surprised faces elicited amygdala responses that were of intermediate magnitude but sustained over time. In this work, amygdala reactivity to unpredictable surprised faces was similar to clearly negative faces, leading us to hypothesize that the unpredictable context led participants to interpret surprised faces more negatively.

To address this hypothesis, we designed a second experiment using facial EMG to measure implicit valence interpretations of ambiguous social cues in unpredictable contexts. Previous research shows that corrugator activity is potentiated by unpleasant emotional pictures and inhibited by pleasant pictures (Cacioppo *et al.*, 1986; Lang *et al.*, 1993), including negative (angry) and positive (happy) facial expressions (Dimberg *et al.*, 2000). Here, we found that corrugator activity was potentiated by unpredictable presentations of surprised faces, similar to clearly negative (angry) expressions, and attenuated by predictable presentations of surprised faces, similar to clearly positive (happy) expressions.

Several findings lend support for this work. First, negative interpretations of neutral faces are associated with greater amygdala activity than positive interpretations (Blasi *et al.,* 2009), and heightened amygdala reactivity to neutral faces is observed in individuals with higher levels of anxiety (Somerville *et al.*, 2004). Our results are consistent with these findings in that we observed an early, relatively enhanced amygdala responses to unpredictable faces, and corrugator responses to these same faces suggest that unpredictability influenced the way that participants responded to the stimuli as though they were interpreting them negatively. However, research has also shown that the amygdala is responsive to clearly positive (i.e. happy) faces (Costafreda *et al.,* 2008; Sergerie *et al.*, 2008). While it is true that the amygdala shows enhanced reactivity to happy faces, the temporal pattern of responses differs from responses to fearful faces. As noted above, fearful faces elicit early high magnitude responses that habituate over time (Kim, 2003; Whalen *et al.,* 2009) whereas happy faces elicit early amygdala responses of an intermediate magnitude that are sustained over time (Somerville *et al.*, 2004). Here we observed an intermediate, sustained amygdala responses to predictable surprised faces, which were also interpreted more positively than unpredictable faces. One remaining question is whether unpredictability could also bias amygdala responses to neutral faces, or to more clearly valenced facial expressions, such as happy or angry faces. We expect that unpredictability might influence interpretations of surprised expressions more than other facial expressions because surprised expressions are emotionally meaningful (high arousal) and are unique in that they can be interpreted positively or negatively, thus rendering them highly susceptible to contextual manipulations (see also Kim *et al.*, 2004). However, future studies including facial expressions that vary along multiple dimensions of ambiguity (e.g. valence, arousal, predictive value) could help clarify the influence of unpredictability on amygdala reactivity to facial expressions in a variety of contexts.

The fact that temporal unpredictability biases amygdala reactivity to social cues is consistent with previous research showing that the amygdala is sensitive to unpredictability itself. For example, Herry *et al.* (2007) showed that amygdala responses to neutral tones were heightened when the tones were presented unpredictably. In addition, a recent review by Diaz-Mataix *et al.* (2014) suggests that the amygdala might play a critical role in processing the timing between conditioned and unconditioned stimulus presentations during Pavlovian conditioning. The amygdala is also sensitive to other forms of uncertainty, such as the predictive value conveyed by facial expressions of emotion (Whalen *et al.*, 2001). The present results build upon these ideas by suggesting that unpredictability can also serve as a context that can help resolve ambiguity.

Limitations

In this report, the EMG and fMRI studies were conducted on different participants. Future studies employing a within subjects design could link individual differences in the magnitude of amygdala reactivity (or rate of habituation) to individual differences in implicit valence interpretations. However, recent research (Heller *et al.*, 2014) simultaneously collecting fMRI and EMG data demonstrated that increased corrugator responses to negative stimuli are associated with increased amygdala responses to those stimuli, suggesting that the patterns of activity reported here might also be observed within participants.

Here, we chose to use a passive viewing paradigm because amygdala responses to emotional stimuli tend to be more robust during passive compared to task-based paradigms (Costafreda *et al.*, 2008). However, such a design is not without limitations. For example, we were able to informally monitor participants' wakefulness during the fMRI study through a camera mounted to the head coil, but have no quantifiable measure of alertness to suggest similar task engagement across studies.

Conclusions

The way that people interpret and respond to facial expressions of emotion plays a critical role in shaping behavior and social interactions, and our interpretations are widely vulnerable to the context in which we encounter these expressions. Here, we manipulated the temporal predictability of surprised facial expressions in order to determine whether unpredictability can bias interpretations of these social cues. We found that neural and physiological responses to these otherwise ambiguous facial expressions became more negative when the faces were unpredictable. These findings add to a growing body of work characterizing the neural circuitry that supports our ability to resolve ambiguity, ultimately shaping the way that we perceive and respond to an often uncertain social environment.

Acknowledgements

The authors would like to thank Tom Johnstone, PhD, Mark Kimak, MS, Rebecca Loucks, BA, Emily C. Mazzulla, PhD, Catherine J. Norris, PhD, Jonathan Oler, PhD, Leah Somerville, PhD and Erika Ruberry, BA. Work supported by National Institutes of Health (MH 069315 and MH 080716). M.J.K. was supported by MH090672.

References

Alvarez, R.P., Chen, G., Bodurka, J., Kaplan, R., Grillon, C. (2011). Phasic and sustained fear in humans elicits distinct patterns of brain activity. *Neuroimage*, **55**, 389–400.

Benjamini, Y. Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society*, **57**, 289–300.

Bensafi, M., Brown, W.M., Tsutsui, T., *et al.* (2003). Sex-steroid derived compounds induce sex-specific effects on autonomic nervous system function in humans. *Behavioral Neuroscience*, **117**, 1125–34.

Blasi, G., Hariri, A.R., Alce, G., *et al.* (2009). Preferential amygdala reactivity to the negative assessment of neutral faces. *Biological Psychiatry*, **66**, 847–53.

Buchel, C., Dolanm R.J., Armony, J.L., Friston, K.J. (1999). Amygdala-hippocampal involvement in human aversive trace conditioning revealed through event-related functional magnetic resonance imaging. *Journal of Neuroscience*, **19**, 10869–76.

Buchel, C., Morris, J., Dolan, R.J., Friston, K.J. (1998). Brain systems mediating aversive conditioning: an event-related fMRI study. *Neuron*, **20**, 947–57.

Bush, L.K., Hess, U., Wolford, G. (1993). Transformations for within-subject designs: a monte carlo investigation. *Psychological Bulletin*, **113**, 566–79.

Cacioppo, J.T., Petty, R.E., Losch, M.E., Kim, H.S. (1986). Electromyographic activity over facial muscle regions can differentiate the valence and intensity of affective reactions. *Journal of Personality and Social Psychology*, **50**, 260–8.

Cacioppo, J.T., Uchino, B.N., Crites, S.L., *et al.* (1992). Relationship between facial expressiveness and sympathetic activation in emotion: a critical review, with emphasis on modeling underlying mechanisms and individual differences. *Journal of Personality and Social Psychology*, **62**, 110–28.

Cheng, D.T., Richards, J., Helmstetter, F.J. (2007). Activity in the human amygdala corresponds to early, rather than late period autonomic responses to a signal for shock. *Learning and Memory*, **14**, 485–90.

Costafreda, S.G., Brammer, M.J., David, A.S., Fu, C.H.Y. (2008). Predictors of amygdala activation during the processing of emotional stimuli: a meta-analysis of 385 PET and fMRI studies. *Brain Research Reviews*, **58**, 57–70.

Davies, C.D., Craske, M.G. (2015). Psychophysiological responses to unpredictable threat: effects of cue and temporal unpredictability. *Emotion*, **15**, 195–200.

Diaz-Mataix, L., Tallot, L., Doyere, V. (2014). The amygdala: a potential player in timing CS-US intervals. *Behavioural Processes*, **101**, 112–22.

Dimberg, U., Thunberg, M., Elmehed, K. (2000). Unconscious facial reactions to emotional facial expressions. *Psychological Science*, **11**, 86–9.

First, M.B., Spitzer, R.L., Gibbon, M., Williams, J.B. (1995). Structured clinical interview for DSM-IV Axis I Disorders (SCID). New York: New York State Psychiatric Institute. *Biometrics Research*.

Fowles, D.C., Christie, M.J., Edelberg, R., Grings, W.W., Lykke, D.T., Venables, P.H. (1981). Publication recommendations for electrodermal measurements. *Psychophysiology*, **18**, 232–9.

Fridlund, A., Cacioppo, J.T. (1986). Guidelines for human electromyographic research. *Psychophysiology*, **23**, 567–89.

Grillon, C., Baas, J.M., Cornwell, B., Johnson, L. (2006). Context conditioning and behavioral avoidance in a virtual reality environment: effect of predictability. *Biological Psychiatry*, **60**, 752–9.

Grillon, C., Baas, J.P., Lissek, S., Smith, K., Milstein, J. (2004). Anxious responses to predictable and unpredictable aversive events. *Behavioral Neuroscience*, **118**, 916–24.

Heller, A.S., Lapate, R.C., Mayer, K.E., Davidson, R.J. (2014). The face of negative affect: trial-by-trial corrugator responses to negative pictures are positively associated with amygdala and negatively associated with ventromedial prefrontal cortex activity. *Journal of Cognitive Neuroscience*, **26**, 2102–10.

Herry, C., Bach, D.R., Esposito, F., *et al.* (2007). Processing of temporal unpredictability in human and animal amygdala. *Journal of Neuroscience*, **27**, 5958–66.

Holland, P.C., Gallagher, M. (1999). Amygdala circuitry in attentional and representational processes. *Trends in Cognitive Science*, **3**, 65–73.

Kapp, B.S., Whalen, P.J., Supple, W.F., Pascoe, J.P. (1992). Amygdaloid contributions to conditioned arousal and sensory information processing. In: Aggleton, J. P., editor. *The Amygdala: Neurobiological Aspects of Emotion, Memory, and Mental Dysfunction*, 229–54, Wiley-Liss, New York, NY.

Kim, H., Somerville, L.H., Johnstone, T., Alexander, A.L., Whalen, P.J. (2003). Inverse amygdala and medial prefrontal cortex responses to surprised faces. *Neuroreport*, **14**, 2317–22.

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

Labar, K.S., Gatenby, J.C., Gore, J.C., Ledoux, J.E., Phelps, E.A. (1998). Human amygdala activation during conditioned fear acquisition and extinction: a mixed-trial fmri study. *Neuron*, **20**, 937–45.

Lang, P.J., Greenwald, M.K., Bradley, M.M., Hamm, A.O. (1993). Looking at pictures: affective, facial, visceral, and behavioral reactions. *Psychophysiology*, **30**, 261–73.

Maldjian, J.A., Laurienti, P.J., Burdette, J.B., Kraft, R.A. (2003). An Automated Method for Neuroanatomic and Cytoarchitectonic Atlas-based Interrogation of fMRI Data Sets. *NeuroImage*, **19**, 1233–9.

Maldjian, J.A., Laurienti, P.J., Burdette, J.H. (2004). Precentral gyrus discrepancy in electronic versions of the Talairach atlas. *Neuroimage*, **21**, 450–5.

Mineka, S., Kihlstrom, J.F. (1978). Unpredictable and uncontrollable events: a new perspective on experimental neurosis. *Journal of Abnormal Psychology*, **87**, 256–71.

Morris, J.S., Buchelm C., Dolan, R.J. (2001). Parallel neural responses in amygdala subregions and sensory cortex during implicit fear conditioning. *Neuroimage*, **13**, 1044–52.

Neta, M., Davis, F.C., Whalen, P.J. (2011). Valence resolution of ambiguous facial expressions using an emotional oddball task. *Emotion*, **11**, 1425–33.

Neta, M., Norris, C.J., Whalen, P.J. (2009). Corrugator muscle responses are associated with individual differences in positivity-negativity bias. *Emotion*, **9**, 640–8.

Norris, C.J., Larsen, J.T., Cacioppo, J.T. (2007). Neuroticism is associated with larger and more prolonged electrodermal responses to emotionally evocative pictures. *Psychophysiology*, **44**, 823–6.

Oldfield, R.C. (1971). The assessment and analysis of handedness: the edinburgh inventory. *Neuropsychologia*, **9**, 97–113.

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**, 865–70.

Pattyn, N., Neyt, X., Henderick, D., Soetens, E. (2008). Psychophysiological investigation of vigilance decrement: Boredom or cognitive fatigue? *Physiology and Behavior*, **93**, 369–78.

Plichta, M.M., Grimm, O., Morgen, K., *et al.* (2014). Amygdala habituation: a reliable fMRI phenotype. *NeuroImage*, **103**, 383–90.

Phelps, E.A., O'connor, K.J., Gatenby, J.C., Gore, J.C., Grillon, C., Davis, M. (2001). Activation of the left amygdala to a cognitive representation of fear. *Nature Neuroscience*, **4**, 437–41.

Robinson, O.J., Letkiewicz, A.M., Overstreet, C., Ernst, M., Grillon, C. (2011). The effect of induced anxiety on cognition: threat of shock enhances aversive processing in healthy individuals. *Cognitive, Affective and Behavioral Neuroscience*, **11**, 217–27.

Sergerie, K., Chochol, C., Armony, J.L. (2008). The role of the amygdala in emotional processing: a quantitative meta-analysis of functional neuroimaging studies. *Neuroscience and Biobehavioral Reviews*, **32**, 811–30.

Shackman, A.J., Maxwell, J.S., McMenamin, B.W., Greischar, L.L., Davidson, R.J. (2011). Stress potentiates early and attenuates late stages of visual processing. *Journal of Neuroscience*, **31**, 1156–61.

Smith, J.C., Bradley, M.M., Lang, P.J. (2005). State anxiety and affective physiology: effects of sustained exposure to affective pictures. *Biological Psychology*, **69**, 247–60.

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.

Tottenham, N., Tanaka, J.W., Leon, A.C., *et al.* (2009). The nimstim set of facial expressions: judgments from untrained research participants. *Psychiatry Research*, **168**, 242–9.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., *et al.* (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*, **15**, 273–89.

Vansteenwegen, D., Iberico, C., Vervliet, B., Marescau, V., Hermans, D. (2008). Contextual fear induced by unpredictability in a human fear conditioning preparation is related to the chronic expectation of a threatening us. *Biological Psychology*, **77**, 39–46.

Weiss, J.M. (1970). Somatic effects of predictable and unpredictable shock. *Psychosomatic Medicine*, **32**, 397–408.

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. (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, J.M., Neta, M. (2009). Human amygdala responses to facial expressions of emotion. In: Phelps, E.A., Whalen P.J., editors. *The Human Amygdala*. New York: Guilford Press, 265–88.

Whalen, P.J., Shin, L.M., McInerney, S.C., Fischer, H., Wright, C.I., Rauch, S.L. (2001). A functional MRI study of human amygdala responses to facial expressions of fear vs. anger. *Emotion*, **1**, 70–83.