The dynamic process of ambiguous emotion perception

Maital Neta , Michael M. Berkebile & Jonathan B. Freeman

Routledge
Taylor & Francis Group

BRIEF ARTICLE

Check for updates

The dynamic process of ambiguous emotion perception

Maital Neta[a], Michael M. Berkebile[b] and Jonathan B. Freeman[b]

[a]Department of Psychology and Center for Brain, Biology, and Behavior, University of Nebraska-Lincoln, Lincoln, NE, USA;
[b]New York University, New York, NY, USA

ABSTRACT

Everyday social interactions hinge on our ability to resolve uncertainty in nonverbal cues. For example, although some facial expressions (e.g. happy, angry) convey a clear affective meaning, others (e.g. surprise) are ambiguous, in that their meaning is determined by the context. Here, we used mouse-tracking to examine the underlying process of resolving uncertainty. Previous work has suggested an initial negativity, in part via faster response times for negative than positive ratings of surprise. We examined valence categorizations of filtered images in order to compare faster (low spatial frequencies; LSF) versus more deliberate processing (high spatial frequencies; HSF). When participants categorised faces as "positive", they first exhibited a partial attraction toward the competing ("negative") response option, and this effect was exacerbated for HSF than LSF faces. Thus, the effect of response conflict due to an initial negativity bias was exaggerated for HSF faces, likely because these images allow for greater deliberation than the LSFs. These results are consistent with the notion that more positive categorizations are characterised by an initial attraction to a default, negative response.

Our daily lives are saturated with affective value (e.g. a visit from a friend, the ringing of an alarm clock, a beautiful sunset, a hot cup of coffee). When we encounter new information (new people, sounds, locations, flavours), we readily sort this information into emotional valence categories: good or bad, approach or avoid, reward or threat. Indeed, the human brain quickly predicts affective value based on previous experiences (i.e. whether something is pleasant/approachable or unpleasant/to-be-avoided; Cabanac, 2002). Facial expressions convey particularly rich information about another person and the environment. Some expressions are clear-cut (angry face predicts threat/avoidance), whereas others are more ambiguous, because they readily predict both rewarding or threatening outcomes. For example, a surprised expression is associated with both positive (a friend's unexpected visit) and negative (hearing that a loved one was in a car accident) outcomes. In other words, when we perceive anger (or happiness)

on another's face, we infer a prototypically negative (or positive) context, whereas when we see a surprised expression, the valence of the context is relatively ambiguous.

Previous research has documented a wide range of individual differences in "valence bias", or the tendency to categorise ambiguous cues (e.g. surprised faces) as having a positive or negative valence (Kim et al., 2003; Neta et al., 2009). Some individuals overwhelmingly tend to categorise such cues as positive, while others overwhelmingly categorise them as negative. Despite these individual differences, Neta and colleagues proposed an initial negativity hypothesis, such that ambiguous cues initially activate a negative valence representation, which can be overridden by a positive representation via an additional mechanism putatively involved in emotion regulation (Petro et al., 2018).

Various behavioural measures and techniques have been used to provide indirect evidence for this

CONTACT Maital Neta ✉ mneta2@unl.edu Center for Brain, Biology, and Behavior, University of Nebraska-Lincoln, B84 East Stadium, Lincoln, NE 68588-0156, USA
Supplemental data for this article can be accessed https://doi.org/10.1080/02699931.2020.1862063.

initial negativity hypothesis. For example, reaction times are longer for positive than negative categorizations of surprised faces (Neta et al., 2009), and encouraging greater deliberation results in more positive categorizations (Neta & Tong, 2016). Experimental manipulations that promote hypervigilance and reduce cognitive control (i.e. stress induction, threat-of-shock), on the other hand, result in more negative categorizations (+Brown et al., 2017). Finally, low spatial frequency (LSF) images emphasising faster visual processing (Bar et al., 2006) result in more negative categorizations of surprised faces than the more elaborate processing of high spatial frequency (HSF) images (Neta & Whalen, 2010). However, while these findings are *suggestive* of an initial negativity (in that, for example, positivity is associated with more processing/response time), there is not clear evidence demonstrating that positive categorizations are characterised by an initial attraction toward a competing negative response.

More recently, research has begun exploiting online hand movements via mouse-tracking to index competing response tendencies during valence categorisation. In mouse-tracking paradigms, participants move the mouse from the bottom-centre of the screen to response options in either top corner (e.g. negative vs. positive). Despite participants' explicit response, mouse trajectories may reveal a simultaneous attraction toward the unselected response option (on the opposite side of the screen). Thus, the paradigm moves beyond a delayed processing/response time for positive categorizations of surprise and directly measures competition from an unselected (negative) response alternative, including its particular millisecond-resolution temporal dynamics (Freeman, 2018). This paradigm therefore provides a sensitive window into the process rather than products of categorisation. One prior study used mouse-tracking to explore valence categorizations of surprised faces, but focused on trajectories that differed as a function of valence bias and cognitive load (Mattek et al., 2016). Thus, rather than examining the dynamic process of valence categorisation across individuals, this work examined trajectories in individuals with a negative versus positive bias and found that bias did not impact trajectories when under high load.

To more directly study the dynamic process of valence categorisation when resolving ambiguity across all individuals, and to provide evidence for the initial negativity hypothesis, we used mouse-tracking to examine response trajectories for LSFs (faster visual processing) and HSFs (more elaborate processing) of facial expressions. We predicted a general negativity bias, such that when participants select the "negative" response their trajectories are especially direct; and when they ultimately select the "positive" response, their trajectories reveal an early bias toward the "negative" response, producing response conflict. Importantly, this response conflict due to the negativity bias should be exacerbated by HSF images, as they promote more deliberative processing that would only serve to intensify the conflict, relative to LSF images. Thus, we should observe negative categorizations of surprise (putatively the default) to be characterised by direct trajectories, but positive categorizations (putatively overriding the initial negativity) to be characterised by indirect trajectories that show response conflict, and this should be pronounced for HSF images.

Method

Participants

A power analysis using G*Power3.1 suggested a total sample size of 101 participants would be necessary to achieve 90% power in detecting significant effects with an effect size comparable to previous work (d = .31; Neta et al., 2009). One-hundred and twenty-five participants were recruited from Amazon Mechanical Turk. Seven participants were excluded because they provided incomplete data (responding on less than 80% of trials), and ten were excluded for failing to accurately rate clearly valenced faces (angry/happy) on at least 60% of trials (as in previous work; Neta & Whalen, 2010). This resulted in a final sample of 106 participants (52 female; ages 18–43 years, M_{age}(SD) = 28.71(4.14)). All procedures were approved by the New York University Committee for the Protection of Human Subjects.

Stimuli

The stimuli were taken from previous work (Neta & Whalen, 2010), including an equal number of male and female faces from NimStim (8 individuals; Tottenham et al., 2009), Pictures of Facial Affect (13 individuals; Ekman & Friesen, 1976), and Averaged Karolinska Directed Emotional Faces (KDEF) databases (39 individuals; Lundqvist et al., 1998). Of interest were responses to surprised expressions; angry and happy

expressions were included as response anchors and to validate performance on the task (Neta et al., 2009). All images were transformed to gray-scale with a resolution of 75 dots per inch. Facial expressions were validated by an independent set of raters labelling each expression; only faces correctly labelled by more than 60% of raters were included.

Further, the original image (broad spatial frequencies, BSFs) was filtered to create two versions of each face: one comprising primarily HSF information and one primarily LSF information (Figure 1). We used a high-pass cutoff of 24 cycles per image for HSFs and a low-pass cutoff of 6 cycles per image for LSFs (Neta & Whalen, 2010). Prior to filtering, we equated contrast and luminance across stimuli. For the task, we used 99 stimuli (face identities posing a particular expression) that were counterbalanced such that each subject viewed a given face as either filtered (counterbalanced order of HSF and LSF) or intact (two

Figure 1. Examples of the stimuli. Broad-spatial frequency (BSF) images were intact. These images were filtered to emphasise the high-spatial frequency (HSF) or low-spatial frequency (LSF) information, in order to promote more elaborate or faster processing, respectively. Each expression was presented either as intact (BSF) or filtered (HSF and LSF) to each participant, and this was counterbalanced across participants.

presentations of the BSF), totalling 198 trials. Faces of all types (BSF, HSF, and LSF) were presented pseudorandomly within each of three runs of 66 trials each. We avoided presenting the same identity in BSF and filtered versions to a given subject so that BSF versions would not affect ratings of the filtered images (Neta & Whalen, 2010).
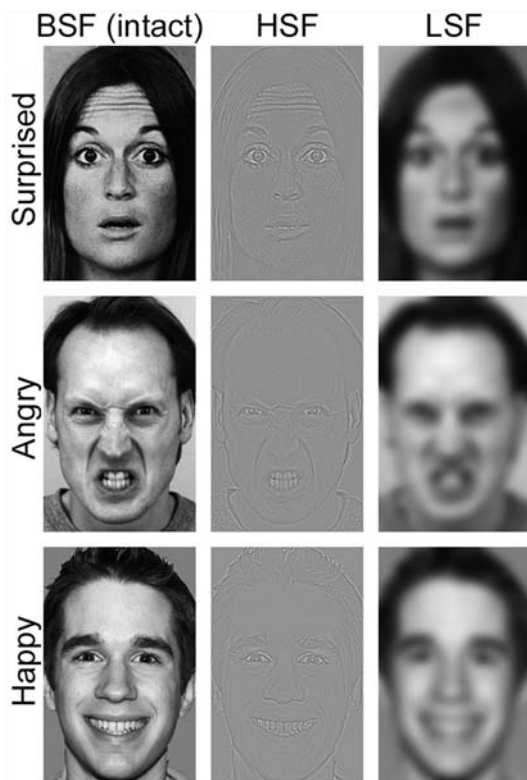
Procedure

Participants categorised faces as positive or negative in a mouse-tracking paradigm using a Javascript implementation of MouseTracker (Freeman & Ambady, 2010). Participants began each trial by clicking "Start" at the bottom centre of the screen, after which a fixation cross appeared for 500 ms, followed by a face for 200 ms. Participants categorised the valence of the face's emotion as quickly and accurately as possible by moving the computer mouse to a "positive" or "negative" response option, which appeared in the upper left and right corners of the screen (counterbalanced across participants), and clicking on the response. There was no time limit within participants needed to complete their response. However, in order to prevent mouse trajectories from being off-line with the decision process, participants were encouraged to start moving within 400 ms of face onset; if they did not, a message appeared following the trial encouraging earlier movement initiation, and the trial was excluded. This criterion resulted in removal of 964 trials out of a total 24,353 across participants (i.e. 3.96% of trials; $M(SD) = 7.84(6.45)$ out of 198 trials per participant). During the categorisation process, the streaming *x, y* coordinates of the mouse were recorded.

Results

Table 1 reports valence ratings scored as the percentage of trials the participant rated as negative out of the total number of trials for that condition (Neta et al., 2009; see Supplementary Material for analysis).

Mouse-trajectory data underwent standard preprocessing (Freeman & Ambady, 2010). All trajectories were rescaled into a standard coordinate space (top-left: "1, 1.5"; bottom-right: "1, 0") and normalised into 100 time bins (101 time-steps) using linear interpolation to permit averaging of their full length across multiple trials. For comparison, all trajectories were remapped rightward. With this orientation, the selected response is located at $x = 1.0$ and unselected

Table 1. Frequency of negativity ratings by condition.

|	Surprised			Angry			Happy		
	BSF	LSF	HSF	BSF	LSF	HSF	BSF	LSF	HSF
Mean (SD)	77.45 (24.59)	78.24 (24.73)	76.83 (23.50)	98.07 (4.29)	94.19 (6.88)	93.69 (9.82)	3.14 (6.94)	4.19 (7.06)	9.48 (15.53)
Range (Min–Max)	0–100	0–100	3.45–100	77.78–100	63.16–100	42.11–100	0–38.89	0–33.33	0–100

Notes: Dependent measure is Percent Negative Ratings, so mean frequency of positive ratings can be calculated by subtracting the mean frequency of negative ratings from 100.

response at $x = -1.0$. Thus, more positive x-coordinates indicated a more direct and facilitated trajectory, while less positive (or even negative) x-coordinates indicated a more indirect trajectory temporarily attracted to the unselected response.

Analyses focused on filtered images of surprised expressions so to emphasise the distinction between faster (LSF) versus more elaborate processing (HSF) of ambiguously valenced stimuli. We used generalised estimating equations multi-level regression (GEE) models, which permits analysis of trial-by-trial data while appropriately accounting for intracorrelations due to the nested data structure (Zeger & Liang, 1986). GEE models of initiation and response times are provided in Supplementary Material. For all GEE models, we report unstandardised regression coefficients (B) and Wald Z-statistics.

Plotting trajectories' x-coordinates over 101 time-steps for the four conditions [Rating (positive, negative) $\times$ Filter (LSF, HSF)] supported our predictions (Figure 2a): trajectories were more direct for negative than positive ratings, and this was exacerbated for HSF images. To probe this Rating $\times$ Filter interaction, at each time-step, we regressed the x-coordinate onto Rating (-0.5 = negative, 0.5 = positive), Filter (-0.5 = LSF, 0.5 = HSF), and the interaction using GEE. To solve the issue of multiple statistical testing at each time-step, we used bootstrapping on 10,000 simulated experiments (see Supplemental Material and Dale et al., 2007 for details). The bootstrapping showed that the experiment-wide significance of the Rating $\times$ Filter interaction was guaranteed at a criterion of $p < .05$, $p < .01$, or $p < .001$ if at least 5, 6, or 8 consecutive time-steps showed a significant effect, respectively.

The Rating $\times$ Filter interaction effect was significant for time-steps 48–57, guaranteeing a significant experiment-wide interaction effect ($p < .001$). To characterise the nature of the interaction, we averaged x-coordinates within these time-steps and submitted them to GEE. There was a significant effect of Rating ($B = -0.11$, SE = 0.03; $Z = -3.88$; $p < .001$), with negative-rating trials showing higher x-coordinates (more direct trajectories) than positive, evidencing a general negativity bias. There was also a significant effect of Filter ($B = -0.06$, SE = 0.01, $Z = -4.74$, $p < .001$), with LSF trials showing higher x-coordinates than HSF trials, as expected given faster processing of LSFs. Notably, as expected given the bootstrapping, these effects were qualified by a significant interaction ($B = -0.06$, SE = 0.03, $Z = -2.10$; $p = .04$), such
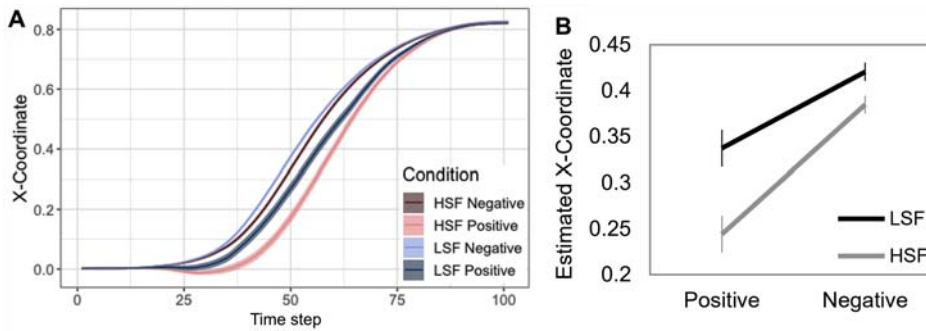
Figure 2. Positive ratings are characterised by an early attraction to negativity particularly in response to HSF faces. (a) There was a partial attraction to the competing (unselected) response option which was more evident on trials rated as positive (attraction to negative) than on trials rated as negative (attraction to positive). This pattern of results was exaggerated for the HSF images, which emphasise a slower, more elaborate processing than the LSF images. (b) We averaged the x-coordinates within the sequence of the trajectory showing a significant interaction (time steps 48–57) and submitted these to GEE multi-level regression. We found that LSF trajectories demonstrated the negativity bias, with significantly more attraction (i.e. lower x-coordinates) toward the opposite response for "positive" rather than "negative" categorizations ($B = -0.08$; SE = 0.03; $Z = -2.52$; $p = .01$), but this effect was exacerbated in the HSF condition ($B = -0.14$, SE = 0.03, $Z = -4.57$, $p < .001$). Error bars denote standard error of the mean.

that LSF trajectories demonstrated a negativity bias (i.e. more attraction toward the competing response for "positive" than "negative" categorizations; $B = -0.08$; SE = 0.03; $Z = -2.52$; $p = .01$) that was exacerbated in the HSF condition ($B = -0.14$, SE = 0.03, $Z = -4.57$, $p < .001$). Together, these results show that, even when participants selected the "positive" response, they exhibited an initial attraction toward the "negative" response overall, and this early "negative" activation was stronger for HSF faces (Figure 2b).

It is possible that what appears to be response conflict and early negativity in the average trajectories could be spuriously produced, for example, if a sub-population of erroneous "flip-flopping" trajectories were averaged together with very direct trajectories (Freeman & Dale, 2013). To rule this out, we inspected the distribution of x-coordinates during time-steps 48–57, as multimodality would suggest multiple sub-populations of trajectories at play (Freeman & Dale, 2013). Hartigan's dip statistic (HDS; Hartigan & Hartigan, 1985) provides an inferential test of multimodality against a null hypothesis of unimodality. There was no evidence of multimodality (LSFs rated as "positive" HDS = .011, $p = .88$ or "negative" HDS = .007, $p = .52$; HSFs rated as "positive" HDS = .011, $p = .931$ or "negative" HDS = .006, $p = .855$).

Another possible explanation is that our results merely reflect perceptual confounds, such that surprised faces show greater resemblance to angry than happy faces. We calculated pixel-based similarity (Pearson correlation between vectorised pixel-maps)

between surprised and angry, and between surprised and happy faces, and compared the correlations (Meng et al., 1992). For HSFs, surprised faces were more similar to happy than angry faces ($p < .001$), contrary to this possibility; for BSFs and LSFs, there was no significant difference (ps = .26 and .12, respectively) with the pattern again biased toward happy faces. Such results favour a genuine negativity bias affecting the categorisation process.

Discussion

We provide direct evidence supporting the initial negativity hypothesis, which posits that the more spontaneous interpretation of ambiguity is negative. Specifically, by using mouse-tracking, which provides a unique window into the dynamic process of categorisation, we demonstrated that negative ratings of ambiguity show more direct response trajectories, but positive ratings are characterised by an attraction to the negative (competing) response option. Notably, filtering images into different spatial frequency bands enabled us to examine trajectories in response to images that receive priority processing (i.e. LSFs are processed first and fast) as compared to those processed more slowly and deliberately (HSFs; Bar et al., 2006). As predicted, the effect of response conflict due to negativity bias was exaggerated for HSFs, likely because these images allow for greater deliberation than the LSFs. Interestingly, this effect was evident despite there being no difference in categorizations of HSF and LSF surprised faces,

suggesting that it was the HSF nature of the images that resulted in increased conflict. Having said that, we note that we did not replicate the finding that LSF surprised faces are rated more negatively than HSFs (Neta & Whalen, 2010). Given that other work using a button press response (i.e. similar to the original report; Park et al., 2016) has established replicability, we propose that response modality may impact this finding.

These findings extend work demonstrating that, despite individual differences in valence bias (Kim et al., 2003; Neta et al., 2009), the initial response is negative (Neta & Whalen, 2010) and positivity arises from greater deliberation (Neta & Tong, 2016). This is also consistent with evidence that the default response to uncertainty is negative (Brosschot et al., 2016). We extend these findings by suggesting that increased deliberation (e.g. longer reaction times) may be associated with greater response conflict. Specifically, we suggest that reaction times are longer for positive than negative trials because positive trials are characterised by a greater attraction to the initial, negative (competing) response option, whereas the negative trials result from more direct trajectories. Notably, this increased response conflict due to negativity bias was exaggerated in response to HSFs, which allow for slower and more deliberate processing; hence, HSFs are more vulnerable to this response conflict. Thus, we leveraged filtered images to more explicitly probe the processes associated with a more direct/automatic (negative) versus more indirect (positive) categorisation.

Recent work suggests that the bottom-up negativity is associated with amygdala activation (Kim et al., 2003), and that a top-down mechanism promotes positivity by overriding the initial negativity using an emotion regulation mechanism akin to cognitive reappraisal (Petro et al., 2018). Cognitive reappraisal is a strategy whereby one reframes or reinterprets an emotionally evocative event in order to change the emotional response (e.g. decreasing negative affect; Gross, 1998). With particular relevance to surprised faces, which have a dual-valence ambiguity, we have proposed that participants override the initial negativity by allowing for a more elaborate (re)interpretation of the expression as positive. The present findings support this notion by demonstrating that a positive interpretation is preceded by an attraction toward a negative rating (i.e. participants are not disengaging from the stimuli, but likely overriding the initial negativity). Importantly, this reinterpretation may result from an implicit form of reappraisal (Braunstein et al., 2017), as individuals are not likely aware of this override process. For this reason, using filtered images here was crucial for disentangling trajectories during faster versus more elaborate processing. Future work could use mouse-tracking to examine processes associated with reappraisal and draw a more direct link to the response competition associated with positive ratings.

Finally, we recently reported that experimental manipulations that promote hypervigilance and reduce cognitive control (e.g. stress) result in more negative categorizations (+Brown et al., 2017). Notably, individuals that were more sensitive to the stressor (greater cortisol reactivity) showed more direct trajectories toward the "negative" response when under stress. Thus, stress was not only associated with greater negativity, but also decreased response competition in the context of negativity. We have also recently demonstrated that individuals that use reappraisal more frequently in their daily lives are more resilient to this stress-related negativity (Raio, Brown & Neta, submitted). Future work will be useful for determining if other simpler methods for promoting elaborate processing (HSF images) might also help to mitigate stress-related negativity.

Several limitations are worth noting. Although we focus on the ambiguity of surprised faces, there is a broader context-dependency of emotional expressions (Aviezer et al., 2012; Barrett, 2014). Indeed, a particular facial configuration is always "ambiguous" to some extent (e.g. Fernández-Dols & Crivelli, 2013). However, our approach stemmed from a social signalling standpoint, suggesting that angry expressions are inferred to be prototypically negative and happy positive, but one is more likely to encounter surprised expressions in a range of contexts (positive and negative). Thus, we do not wish to imply a fundamental distinction between angry/happy and surprised expressions' potential for ambiguity, and future work could test how our findings generalise to cases in which angry/happy expressions are also considered relatively ambiguous. Also, although this paradigm characterises individual differences in negativity bias, faces are almost never seen without context in the real world. However, even then, the context is not always sufficient for determining a clear emotional valence (e.g. seeing people in a dark alley either in search of a night club or someone to rob). Our valence bias task intentionally withholds context so that individuals must resolve the

uncertainty by relying on their own biases. Having said that, it could be that this paradigm is a better measure of one's tendency to infer that *situations* are negative or positive, and we are generally agnostic to this distinction; the important feature here is that we can identify one's bias to perceive a more negative or positive meaning in circumstances where both alternatives are valid with relatively equal probability. Future research could explore these open questions, for example, using paradigms lacking categorical emotion labels.

Conclusions

By leveraging mouse-tracking technology and filtered images that emphasise faster versus more deliberate processing, this work provides direct evidence in support of the initial negativity hypothesis. Specifically, we demonstrate the presence of an early negativity bias that weighs in during the process of resolving emotional ambiguity, even when participants ultimately arrive at a "positive" evaluation. This work lends insight into the underlying mechanism of classifying new, and particularly ambiguous information into emotional valence categories.

Acknowledgements

Disclosure statement

Funding

Data availability

All materials, data, and scripts are available at https://osf.io/dar7m/.

References

Aviezer, H., Trope, Y., & Todorov, A. (2012). Holistic person processing: Faces with bodies tell the whole story. *Journal of Personality and Social Psychology*, *103*(1), 20. https://doi.org/10.1037/a0027411

Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmid, A. M., Dale, A. M., Hämäläinen, M. S., Marinkovic, K., Schacter, D. L., Rosen, B. R., & Halgren, E. (2006). Top-down facilitation of visual recognition. *Proceedings of the National Academy of Sciences, USA*, *103*(2), 449–454. https://doi.org/10.1073/pnas.0507062103

Barrett, L. F. (2014). The conceptual act theory: A précis. *Emotion Review*, *6*(4), 292–297. https://doi.org/10.1177/1754073914534479

Braunstein, L. M., Gross, J. J., & Ochsner, K. N. (2017). Explicit and implicit emotion regulation: A multi-level framework. *Social Cognitive and Affective Neuroscience*, *12*(10), 1545–1557. https://doi.org/10.1093/scan/nsx096

Brosschot, J. F., Verkuil, B., & Thayer, J. F. (2016). The default response to uncertainty and the importance of perceived safety in anxiety and stress: An evolution-theoretical perspective. *Journal of Anxiety Disorders*, *41*, 22–34. https://doi.org/10.1016/j.janxdis.2016.04.012

+Brown, C. C., +Raio, C. M., & Neta, M. (2017). Cortisol response enhance negative valence perception for ambiguous facial expressions. *Nature: Scientific Reports*, *7*, article number: 15107. https://doi.org/10.1038/s41598-017-14846-3. (+co-first authors).

Cabanac, M. (2002). What is emotion? *Behavioural Processes*, *60*(2), 69–83. https://doi.org/10.1016/S0376-6357(02)00078-5

Dale, R., Kehoe, C., & Spivey, M. J. (2007). Graded motor responses in the time course of categorizing atypical exemplars. *Memory and cognition*, *35*(1), 15–28.

Ekman, P., & Friesen, W. V. (1976). *Pictures of facial affect* [Slides]. Consulting Psychologists Press.

Fernández-Dols, J. M., & Crivelli, C. (2013). Emotion and expression: Naturalistic studies. *Emotion Review*, *5*(1), 24–29. https://doi.org/10.1177/1754073912457229

Freeman, J. B. (2018). Doing psychological science by hand. *Current Directions in Psychological Science*, *27*(5), 315–323. https://doi.org/10.1177/0963721417746793

Freeman, J. B., & Ambady, N. (2010). Mousetracker: Software for studying real-time mental processing using a computer mouse-tracking method. *Behavior Research Methods*, *42*(1), 226–241. https://doi.org/10.3758/BRM.42.1.226

Freeman, J. B., & Dale, R. (2013). Assessing bimodality to detect the presence of a dual cognitive process. *Behavior Research Methods*, *45*(1), 83–97. https://doi.org/10.3758/s13428-012-0225-x

Gross, J. J. (1998). Antecedent-and response-focused emotion regulation: Divergent consequences for experience, expression, and physiology. *Journal of Personality and Social Psychology*, *74*(1), 224. https://doi.org/10.1037/0022-3514.74.1.224

Hartigan, J. A., & Hartigan, P. M. (1985). The dip test of unimodality. *The Annals of Statistics*, *13*(1), 70–84. https://doi.org/10.1214/aos/1176346577

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*(18), 2317–2322. https://doi.org/10.1097/00001756-200312190-00006

Lundqvist, D., Flykt, A., & Öhman, A. (1998). *The Karolinska directed emotional faces—KDEF* [CD ROM]. Department of

Clinical Neuroscience, Psychology Section, Karo- linska Institutet.

Mattek, A. M., Whalen, P. J., Berkowitz, J. L., & Freeman, J. B. (2016). Differential effects of cognitive load on subjective versus motor responses to ambiguously valenced facial expressions. *Emotion*, *16*(6), 929. https://doi.org/10.1037/emo0000148

Meng, X. L., Rosenthal, R., & Rubin, D. B. (1992). Comparing correlated correlation coefficients. *Psychological Bulletin*, *111*(1), 172. https://doi.org/10.1037/0033-2909.111.1.172

Neta, M., Norris, C. J., & Whalen, P. J. (2009). Corrugator muscle responses to surprised facial expressions are associated with individual differences in positivity-negativity bias. *Emotion*, *9*(5), 640–648. https://doi.org/10.1037/a0016819

Neta, M., & Tong, T. T. (2016). Don't like what you see? Give it time: Longer reaction times associated with increased positive affect. *Emotion*, *16*(5), 730–739. https://doi.org/10.1037/emo0000181

Neta, M., & Whalen, P. J. (2010). The primacy of negative interpretations when resolving the valence of ambiguous facial expressions. *Psychological Science*, *21*(7), 901–907. https://doi.org/10.1177/0956797610373934

Park, G., Vasey, M. W., Kim, G., Hu, D. D., & Thayer, J. F. (2016). Trait anxiety is associated with negative interpretations when resolving valence ambiguity of surprised faces. *Frontiers in Psychology*, *7*, 1164. https://doi.org/10.3389/fpsyg.2016.01164

Petro, N. M., Tong, T. T., Henley, D. J., & Neta, M. (2018). Individual differences in valence bias: fMRI evidence of the initial negativity hypothesis. *Social, Cognitive, and Affective Neuroscience*, *13*(7), 687–698. https://doi.org/10.1093/scan/nsy049

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., Marcus, D. J., Westerlund, A., Casey, B. J., & Nelson, C. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research*, *168*(3), 242–249. https://doi.org/10.1016/j.psychres.2008.05.006

Zeger, S. L., & Liang, K. Y. (1986). Longitudinal data analysis for discrete and continuous outcomes. *Biometrics*, 121–130. https://doi.org/10.2307/2531248