Acta Psychologica 180 (2017) 155–159

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy

Blame everyone: Error-related devaluation in Eriksen flanker task☆

Andrey Chetverikov[a,b,c,*], Polina Iamschinina[c], Alena Begler[c], Ivan Ivanchei[b,c], Margarita Filippova[c], Maria Kuvaldina[c]

[a] Laboratory for Visual Perception and Visuomotor Control, Faculty of Psychology, School of Health Sciences, University of Iceland, Reykjavik, Iceland
[b] Cognitive Research Lab, Russian Academy of National Economy and Public Administration, Moscow, Russia
[c] Department of Psychology, Saint Petersburg State University, Saint Petersburg, Russia

ARTICLE INFO

Keywords:
Error-related devaluation
Errors
Conflict
Negative affect
Flanker

ABSTRACT

Preferences are determined not only by stimuli themselves but also by the way they are processed in the brain. The efficacy of cognitive processing during previous interactions with stimuli is particularly important. When observers make errors in simple tasks such as visual search, recognition, or categorization, they later dislike the stimuli associated with errors. Here we test whether this error-related devaluation exists in Erisken flanker task and whether it depends on the distribution of attention. We found that both attended stimuli (targets) and ignored ones (distractors) are devaluated after errors on compatible trials but not incompatible ones. The extent of devaluation is similar for targets and distractors, indicating that distribution of attention does not significantly influence the attribution of error-related negative affect. We discuss this finding in light of the possible mechanisms of error-related devaluation.

1. Introduction

A softness of touch, a pleasant taste, or an elegant shape – all these qualities could be legitimate reasons for preferring one thing over the other. Yet, previous studies show that preferences depend on cognitive processing as much as on the intrinsic qualities of stimuli (Albrecht & Carbon, 2014; Chetverikov & Kristjánsson, 2016; Muth & Carbon, 2013; Reber, Schwarz, & Winkielman, 2004; Van de Cruys & Wagemans, 2011). The efficacy of cognitive processing is particularly important: errors result in a negative affect and devaluation of stimuli associated with errors even when participants do not receive any feedback about their accuracy (Aarts, De Houwer, & Pourtois, 2012; Chetverikov, 2014; Chetverikov & Filippova, 2014; Schouppe et al., 2014). Physiological studies also show that activation of reward-related brain regions, such as ventral striatum, depends on response accuracy even when no external feedback is provided (Daniel & Pollmann, 2014; Satterthwaite et al., 2012). Similarly, a fast error-related response-locked negative deflection of brain electrical activity known as error-related negativity (ERN) consistently correlates with negative affect (Aarts, De Houwer, & Pourtois, 2013; Hajcak, McDonald, & Simons, 2004; Luu, Collins, & Tucker, 2000; Moser, Moran, Schroder, Donnellan, & Yeung, 2013; Schroder, Moran, Infantolino, & Moser, 2013). One possible interpretation of this phenomenon is that "marking" error-related stimuli with negative affect might help guide future behavior (Chetverikov & Kristjánsson, 2016). In real life, however, there is usually more than one object present at a time. It is not clear how the negative affect resulting from an error becomes associated with a particular stimulus. Filling this gap is important to understand better both how the preferences are formed in general and how people learn from their errors.

In previous studies of error-related negative affect, usually, only a single stimulus was presented on the screen when an error occurred. For example, Chetverikov (2014) demonstrated that preferences towards previously shown stimuli depend on whether or not observers recognize these stimuli in an unexpected recognition test before preferences were rated. This in sharp contrast to a well-known mere exposure effect suggesting that previously seen stimuli are preferred to novel ones even when they were not consciously perceived (Bornstein, 1989; Zajonc, 1980, 2001). In a meta-analysis of previous studies and several new experiments, Chetverikov (2014) found a typical mere exposure effect only when observers recognized the stimuli. But in case of recognition failure, that is, when observers erroneously thought that the stimulus they are asked to recognize was not presented before, the preferences became more negative as the number of previous exposures increased. This phenomenon was coined error-related devaluation: recognition error results in negative affect that counteracts positive

☆ The studies reported in this paper were supported by the Russian Foundation for Basic Research (research project #15-06-07417a). The authors are grateful to A. Fedorova for assistance with the data collection. The data and scripts for analyses reported in this paper are available at https://osf.io/mtza2/.
* Corresponding author at: Department of Psychology, Saint Petersburg State University, nab. Makarova 6, Saint Petersburg 199034, Russia.
E-mail address: andrey@hi.is (A. Chetverikov).

http://dx.doi.org/10.1016/j.actpsy.2017.09.008
Received 6 September 2016; Received in revised form 8 September 2017; Accepted 18 September 2017
0001-6918/

effects of mere exposure. Later, similar negative effects of errors on preferences were found in visual search (Chetverikov, Jóhannesson, & Kristjánsson, 2015) and categorization tasks (Chetverikov & Filippova, 2014).

The main question of the present study is how error-related devaluation is distributed between stimuli present at the moment of error. In the real world, observers always perceive more than one stimuli. How do they determine which one is to “blame” for the error? Studies of affective misattribution (Payne & Lundberg, 2014; Schwarz & Clore, 1983) demonstrate that affect can automatically spread from one stimulus to another when they are close in time. Then, error-related devaluation might be not limited to stimuli evoking the errors. In support of this hypothesis, Aarts et al. (2012) found that false alarms in a Go/NoGo task speed up subsequent evaluative categorization of negative words compared to positive words. Using similar evaluative categorization procedure to measure affect, Schouppe et al. (2014) found that after errors in Eriksen flanker task (Eriksen & Eriksen, 1974) observers tend to categorize the subsequently presented words as negative more often. These findings indicate that error-related negative affect might diffuse from one stimulus to another. Notably, in these studies both error-related stimuli and subsequently presented words are attended. However, Chetverikov et al. (2015) found that in the visual search task errors do not affect the evaluation of distractors. While liking ratings of the targets became more positive with an increase in search time on correct trials and more negative on error trials, for distractors search time was positively correlated with liking independent of trial accuracy.

We hypothesized that attention might play an important role in error-related devaluation such that only attended stimuli are devaluated. To test this hypothesis, in the present study we conducted an experiment utilizing a modified Eriksen flanker task. In the flanker task, observers had to make decisions about the stimulus presented in the center (target) while surrounding stimuli (distractors, or flankers) were to be ignored. We expected that targets would be devaluated more than distractors following incorrect responses due to the distribution of attention.

In addition, we wanted to test if response accuracy would interact with the trial compatibility. Chetverikov and Kristjansson (2016) suggested that error-related devaluation can result from inconsistency between predictions based on a variety of cues involved in decision-making process. Each decision can utilize different cues: recognition, for example, can be based on shape, colour, semantics, and many other aspects of stimuli. Monitoring this consistency can then help to monitor response accuracy even in the absence of external feedback. Similar ideas were proposed within conflict-monitoring theory (Botvinick, Braver, Barch, Carter, & Cohen, 2001; Yeung, Botvinick, & Cohen, 2004; Yeung & Summerfield, 2012) and self-consistency model of confidence (Koriat, 2011, 2012). In support of this idea, previous studies indicate that the amount of information available for correct responses correlates with the post-error devaluation. For example, longer gaze times on target stimuli in visual search (Chetverikov et al., 2015) or more exposure (Chetverikov, 2014) result in more pronounced post-error devaluation. In the flanker task, compatible trials provide more cues for a correct response than incompatible ones and hence error-related devaluation also should be stronger in the former case than in the latter.

Affective responses to stimuli in a flanker task were studied before by Martiny-Huenger, Gollwitzer, and Oettingen (2014). They found that distractors used in incompatible trials were disliked compared to targets or novel stimuli. Targets, however, were rated similarly regardless of trial compatibility. A subsequent recognition test did not indicate that observers remember the stimuli from the flanker task despite the fact that distractors were devaluated. However, Martiny-Huenger et al. (2014) did not analyse the response accuracy. Thus, while their study provides data regarding the effect of compatibility on preferences, it does not help understand how observers associate error-related negative affect with particular stimuli. Answering this question will reveal the mechanisms of error-related devaluation and the involvement of attention in this process. In the present study, we fill this gap and describe the preferences towards targets and distractors as a function of trial compatibility and response accuracy.

2. Method

2.1. Participants

Sixty-one observer (44 women, 18–31 years old, age $Mdn = 21$) at Saint Petersburg State University voluntarily participated. They were not paid for participation. All reported normal or corrected-to-normal visual acuity. Three participants were excluded because of very long response times on evaluation trials (5.6, 8.8, and 10.3 s as compared to the average of 1.7 s).

2.2. Materials

The experiment was run using PsychoPy 1.81.02 (Peirce, 2007, 2009). Observers sat at approximately 50 cm distance from a 17 in. LCD display with 1280 × 1024 resolution (LG Flatron L1718S). Both target and distractors in the flanker task were grayscale female or male faces tinted with 50% transparent green or blue colours ([0, 1, 0] or [0, 0, 1] in − 1 to 1 RGB colour space). For each observer, twenty-four target-distractor pairs were chosen randomly from a set of 32 male and 32 female faces obtained from Facial Recognition Technology database[1] (Phillips, Moon, Rizvi, & Rauss, 2000; Phillips, Wechsler, Huang, & Rauss, 1998). The same stimuli without tint were used in the subsequent preference task. For the training session, a different set of 40 faces randomly selected for each observer from the same database were used.

2.3. Procedure

The experiment was split into two blocks.[2] In each block observers first completed flanker task and then evaluated the stimuli (Fig. 1). In the flanker task on each trial first the fixation cross was shown for 500 ms. Then a target (in the centre) and four identical distractors (on each side of target) were shown. Response time was limited to 600 ms. Response time was limited to ensure that there will be enough errors for analyses. If observers did not respond within the allocated time, a feedback “TOO LATE” appeared for 500 ms after the response (this response was not included in the following analyses). The stimuli were either 2 or 3° of visual angle (v.a.). On each trial, stimuli sizes were selected randomly to increase the probability of object-based or feature-based inhibition instead of location-based inhibition. Centre-to-centre distance between target and flankers was either 2.5 or 3.8° for smaller and larger stimuli, respectively. Distance depended on size – larger distance (3.8°) was used for larger stimuli. The observers had to determine the colour of a centrally presented face while ignoring the rest of the stimuli by pressing ‘A’ or ‘D’ key marked with green or blue colours, accordingly.

Twenty-four target-distractor pairs were repeated five times each resulting in a total of 120 trials. On compatible trials, target and distractors had the same tint (irrespective of their gender) while on incompatible trials the colours of target and distractors were different. Trial compatibility, target colour, and target gender were counterbalanced.

In the second part of the block on each of 24 trials observers were

[1] Portions of the research in this paper use the FERET database of facial images collected under the FERET programme, sponsored by the DOD Counterdrug Technology Development Program Office.

[2] We have tested for the effect of Block along with its interactions in the analyses reported in this paper but neither Block nor its interactions were significant. The inclusion of Block in analysis also did not affect any conclusions regarding the other effects. Thus, the models reported here do not include Block.

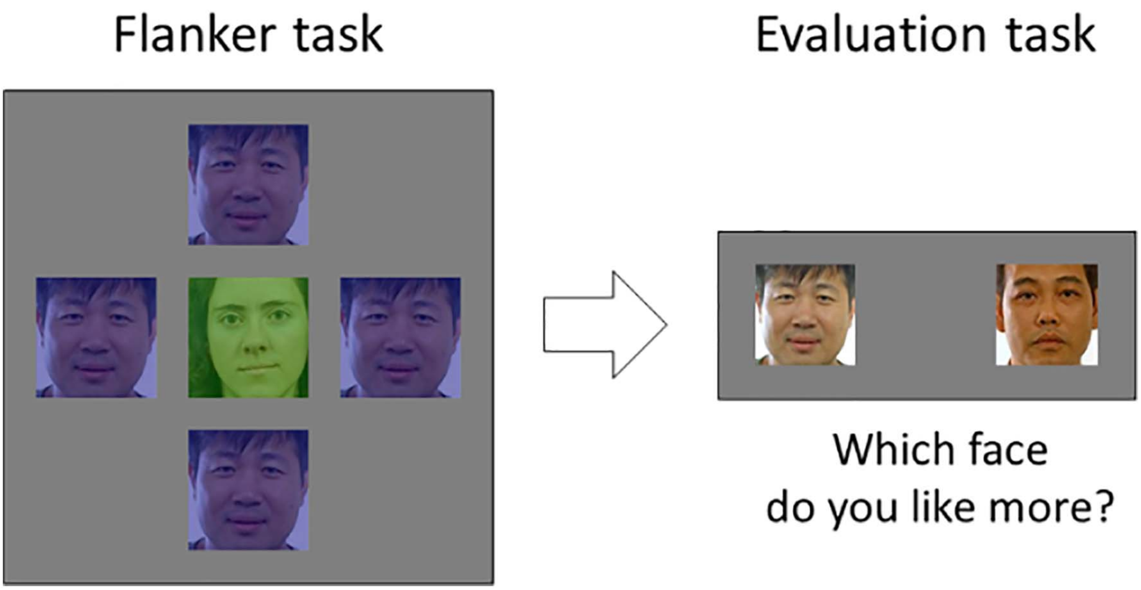

Fig. 1. Stimuli and design. In each of two blocks, observers first performed a flanker task in which they had to determine the colour of a centrally presented face while ignoring the flankers (120 trials per block). Then, observers were presented with pairs of faces (targets or distractors from the flanker task) and were asked to choose the face they liked more (24 trials per block).

first shown a fixation point for 500 ms. Then after a pause of 500 ms they were asked to choose among the two faces the one they like more. Each pair consisted of a target and distractor from the flanker task. Stimuli were resized to 2.5° of v.a. and shown 3° to the left or to the right from screen centre. Target position was counterbalanced. Participants responded by pressing “left” or “right” key, response time was not limited. The faces were presented without tint. Targets and distractors were intermixed randomly so that 25% of the pairs consisted of target and distractors that were on compatible trials, 25% had targets from compatible trials and distractors from incompatible trials, 25% had targets from incompatible trials and distractors from compatible trials, and the last 25% had both targets and distractors from incompatible trials. This way it was possible to estimate separately the effects of compatibility for both targets (further referred to as Target Compatibility) and distractors (Distractor Compatibility). Note that with this design the effects of the presentation itself are balanced because both targets and distractors have the same presentation duration.

Before the main part of the study, observers took part in a training session to familiarize themselves with the flanker task. The training session was immediately before the main part of the experiment. It consisted of series of 20 flanker trials with the procedure described above and was repeated until observers reached 75% accuracy and no more than 10% of “too late” trials.

3. Results

3.1. Flanker effect

Observers' accuracy was lower on incompatible trials, than on compatible ones, $F(1, 57) = 61.33$, $p < .001$, $\eta^2_G = 0.17$ (Table 1).

Table 1
The share of accurate answers, errors, and “too late” responses and the response times for correct responses in the flanker task.

	Correct		Error		“Too late”		RT (correct trials, ms)	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Compatible	0.87	0.07	0.08	0.06	0.06	0.04	421	19
Incompatible	0.80	0.09	0.13	0.08	0.08	0.04	441	18

Correct responses were slower on incompatible trials than on compatible ones, $F(1, 57) = 156.88$, $p < .001$, $\eta^2_G = 0.23$.

3.2. Preferences

To analyse the preferences, we used mixed-effects binomial regression with *lme4* package in R (Bates, Maechler, Bolker, & Walker, 2014) with a choice in the preference task (target preferred over distractors or vice versa) as a dependent variable. Mixed-effects regression allowed to control for the additional noise associated with differences in attractiveness of targets and distractors simultaneously.

We started with a simple model including two main fixed effects: Target Compatibility and Distractor Compatibility, along with their interaction.[3] The random-effects structure was built iteratively (Bates, Kliegl, Vasishth, & Baayen, 2015). Initially, we included three groups of random effects, one for subjects, target stimuli, and distractor stimuli. For each group we then fitted the random effects corresponding to the included fixed effects, starting with the maximal model (including all possible random effects) and removing non-significant random effects one by one. The resulting model included only the random intercepts for target and distractor stimuli along with the main effects of Target Compatibility, Distractor Compatibility, and their interaction.

Overall, observers did not show preference towards targets or distractors (the probability of choosing target over distractor did not differ from chance level, $B = 0.03$, $SE = 0.10$, $Z = 0.29$, $p = .774$). Neither the distractor compatibility, $B = 0.01$, $SE = 0.08$, $Z = 0.12$, $p = .901$, nor the target compatibility, $B = 0.07$, $SE = 0.08$, $Z = 0.79$, $p = .427$, nor their interaction, $B = 0.04$, $SE = 0.16$, $Z = 0.24$, $p = .809$, were significant.

We then proceeded to analyse the effect of response accuracy in the flanker task on preferences. Given that the average number of errors was relatively small despite the time pressure (Table 1), we compared the preferences for stimuli used in trials during which no errors were made through all five repetitions with the stimuli used in trials that resulted in at least one error. Two terms, response accuracy for targets

[3] There could be no main effect of stimulus type, because preferring a distractor would simultaneously mean not choosing a target. In other words, if targets and distractors were compared to novel items, we would be able to construct a Stimulus Type × Compatibility matrix for predictors. But because they are affecting same choice, doing so would result in artificial use of the same trial result twice.

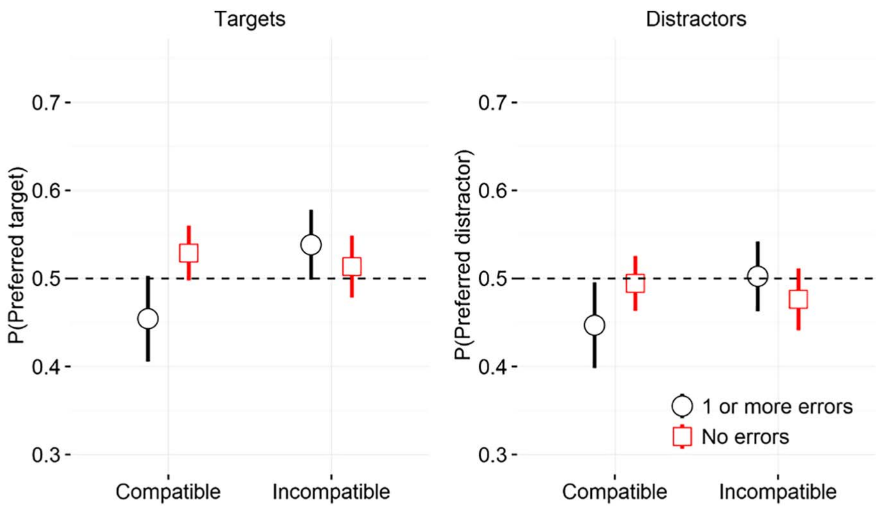

Fig. 2. Preferences towards the targets and the distractors as a function of response accuracy in the flanker task. Bars show 95% confidence intervals (CI). The horizontal line shows the chance level. The probability of preferences above chance level indicates positive evaluation while those below chance level indicate negative evaluation.

and for distractors, were added to the model described above along with their interactions with compatibility for targets and distractors, respectively. Mixed-effect regression indicated a tendency-level main effects of target compatibility, $B = -0.08$, $SE = 0.04$, $Z = 1.78$, $p = .075$, target response accuracy, $B = -0.08$, $SE = 0.04$, $Z = 1.93$, $p = .054$, and a significant interaction between them, $B = -0.12$, $SE = 0.04$, $Z = 2.72$, $p = .007$. For distractors, both main effects were not significant, but their interaction was marginally significant, $B = 0.08$, $SE = 0.04$, $Z = 1.92$, $p = .055$. Fig. 2 demonstrates that the effect of response accuracy was similar for targets and distractors. Post-hoc comparisons between correct and incorrect trials showed that both targets ($B = 0.41$, $SE = 0.13$, $p = .002$) and distractors ($B = -0.27$ $SE = 0.13$, $p = .039$) were rated more positively following correct responses than following errors on compatible trials. On incompatible trials, however, no such difference was observed ($B = -0.07$ $SE = 0.12$, $p = .565$ for targets and $B = 0.07$, $SE = 0.12$, $p = .572$ for distractors).

As shown above, the interaction effect between response accuracy and trial compatibility was numerically more pronounced for targets than for distractors. To test whether this difference was significant, we compared resulting regression model with a restricted model in which the regression coefficients for the two interactions were assumed to be equal. The results showed that the difference in regression coefficients is not significant, $\chi^2(1) = 0.35$, $p = .556$. Similarly, when we analyzed targets and distractors from compatible trials only, devaluation effects for targets, $B = -0.30$, $SE = 0.09$, $Z = 3.18$, $p = .001$, and distractors, $B = 0.23$, $SE = 0.09$, $Z = 2.41$, $p = .016$, were not different from each other, $\chi^2(1) = 0.40$, $p = .527$.

4. Discussion

The results of this study demonstrate error-related devaluation in Eriksen flanker task: both targets and distractors were less likely to be preferred if there was an error associated with these stimuli during preceding flanker trials. We thus extend the scope of the devaluation effect previously found for recognition (Chetverikov, 2014), visual search (Chetverikov et al., 2015), and categorization (Chetverikov & Filippova, 2014). The present study utilized comparisons between similarly exposed stimuli, ruling out the mere exposure explanation. We also controlled for the stimuli effect by including random effects for targets and distractors in mixed-effects regression. Hence, the results cannot be explained by inherent differences between stimuli.

We show for the first time that devaluation occurs not only for target items but for distractors as well. The strength of the effect was similar for target and distractors. This result is important for understanding the mechanisms of error-related devaluation. It shows that when observers made an error, they automatically associated error-induced negative affect with the stimuli at hand. It is particularly interesting that devaluation is observed on compatible trials in which distractors are not detrimental to performance. Though we did not include neutral stimuli in our study it is widely known that compatible trials facilitate performance (e.g., Lamers & Roelofs, 2011). If stimuli are devaluated in order to be avoided in future, then devaluation of distractors may not be functional in a sense that they are not what causes an error and there is no logical reason to avoid them in future. On the other hand, if attribution of negative affect to its source requires additional attentional resources or if observers employ “better safe than sorry” strategy, then it might be more optimal to devaluate all the items associated with the error. That is what we observed in the present study.

This result is consistent with other studies showing that affect can be misattributed to the stimuli close in temporal and spatial context to its source (Payne & Lundberg, 2014; Schwarz & Clore, 1983). As shown by Schouppe et al. (2014), observers tend to categorize words following errors as negative more often. This result can be seen as a devaluation as well though it might also reflect other processes such as higher activation of semantic network related to negative connotations of the categorized words. In our study, the devaluation is clearly at work: errors and evaluation task were temporally separated and hence more negative ratings reflect changes in the evaluation of stimuli rather than momentary negative affect.

Interestingly, Chetverikov et al. (2015) did not find error-related devaluation for distractors in a visual search task. Rather, distractors were unaffected by response accuracy and more exposure (measured by the gaze dwell time on stimuli) was associated with more positive ratings. There are several differences between the two paradigms that can explain the result. First, in our study exposure times were shorter because flanker task is easier than visual search. In addition, in our study errors and evaluation task were separated in time while in the previously reported visual search experiment evaluation occurred immediately after each trial. Thus, it is possible that positive effect from mere exposure is stronger immediately after the trial and might have obscured negative effect from errors in the previously reported results. However, we believe that this explanation is unlikely because exposure time was controlled by Chetverikov et al. (2015) through analyses of gaze dwell time. Second, in the visual search study observers had to find a target face among nine faces by conjunction of tint and gender. This is a more difficult task that requires more elaborated analysis of each stimulus than the flanker task used in the current study. It is possible that more elaborated processing led to more restricted attribution of the negative affect following errors. We believe that the latter explanation is more likely; however, more studies are necessary before drawing conclusions.

In addition, we found that the devaluation effect occurs when errors are committed in the compatible but not the incompatible trials. Stimuli were distributed randomly between conditions, again indicating that stimuli qualities cannot explain the devaluation. According to the affective feedback model (Chetverikov & Kristjánsson, 2016) error-related devaluation is stronger when there is more evidence in favor of correct response. On compatible trials, observers have stronger evidence in favor of correct answer. If they nevertheless make an error there is less consistency between different cues involved in the decision (see also Koriat, 2012; Yeung et al., 2004), and the error-related devaluation should be more pronounced. The results are consistent with this prediction.

In sum, the present study shows that both attended and unattended stimuli are devaluated in flanker task. Instead of attributing negative affect to targets, observers "blame" every stimulus that is present when they make an error. This finding shows that even if one is not engaged in the processing of a specific object, a simple coincidence with an error might be enough for a relatively stable dislike towards this object. Given the amount of cognitive processing that we engage in at every moment of our life, one can only wonder if any of the things that surround us are unaffected by such errors.

References

Aarts, K., De Houwer, J., & Pourtois, G. (2012). Evidence for the automatic evaluation of self-generated actions. *Cognition, 124*(2), 117–127. http://dx.doi.org/10.1016/j.cognition.2012.05.009.

Aarts, K., De Houwer, J., & Pourtois, G. (2013). Erroneous and correct actions have a different affective valence: Evidence from ERPs. *Emotion (Washington, D.C.), 13*(5), 960–973. http://dx.doi.org/10.1037/a0032808.

Albrecht, S., & Carbon, C.-C. (2014). The fluency amplification model: Fluent stimuli show more intense but not evidently more positive evaluations. *Acta Psychologica, 148*(951), 195–203. http://dx.doi.org/10.1016/j.actpsy.2014.02.002.

Bates, D. M., Kliegl, R., Vasishth, S., & Baayen, H. (2015). *Parsimonious Mixed Models Douglas Bates.* 1–27 Retrieved from. https://arxiv.org/abs/1506.04967.

Bates, D. M., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. Retrieved from https://cran.r-project.org/package=lme4.

Bornstein, R. F. (1989). Exposure and affect: Overview and meta-analysis of research, 1968–1987. *Psychological Bulletin, 106*(2), 265–289. http://dx.doi.org/10.1037/0033-2909.106.2.265.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108*(3), 624–652 Retrieved from. https://www.ncbi.nlm.nih.gov/pubmed/11488380.

Chetverikov, A. (2014). Warmth of familiarity and chill of error: Affective consequences of recognition decisions. *Cognition and Emotion, 28*(3), 385–415. http://dx.doi.org/10.1080/02699931.2013.833085.

Chetverikov, A., & Filippova, M. (2014). How to tell a wife from a hat: Affective feedback in perceptual categorization. *Acta Psychologica, 151*, 206–213. http://dx.doi.org/10.1016/j.actpsy.2014.06.012.

Chetverikov, A., Jóhannesson, Ó. I., & Kristjánsson, Á. (2015). Blaming the victims of your own mistakes: How visual search accuracy influences evaluation of stimuli. *Cognition and Emotion, 29*(6), 1091–1106. http://dx.doi.org/10.1080/02699931.2014.968097.

Chetverikov, A., & Kristjánsson, Á. (2016). On the joys of perceiving: Affect as feedback for perceptual predictions. *Acta Psychologica, 169*, 1–10. http://dx.doi.org/10.1016/j.actpsy.2016.05.005.

Daniel, R., & Pollmann, S. (2014). A universal role of the ventral striatum in reward-based learning: Evidence from human studies. *Neurobiology of Learning and Memory, 114*, 90–100. http://dx.doi.org/10.1016/j.nlm.2014.05.002.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics, 16*(1), 143–149. http://dx.doi.org/10.3758/BF03203267.

Hajcak, G., McDonald, N., & Simons, R. F. (2004). Error-related psychophysiology and negative affect. *Brain and Cognition, 56*(2), 189–197. http://dx.doi.org/10.1016/j.bandc.2003.11.001.

Koriat, A. (2011). Subjective confidence in perceptual judgments: A test of the self-consistency model. *Journal of Experimental Psychology: General, 140*(1), 117–139. http://dx.doi.org/10.1037/a0022171.

Koriat, A. (2012). The self-consistency model of subjective confidence. *Psychological Review, 119*(1), 80–113. http://dx.doi.org/10.1037/a0025648.

Lamers, M. J. M., & Roelofs, A. (2011). Attentional control adjustments in Eriksen and Stroop task performance can be independent of response conflict. *The Quarterly Journal of Experimental Psychology, 64*(6), 1056–1081. http://dx.doi.org/10.1080/17470218.2010.523792.

Luu, P., Collins, P., & Tucker, D. M. (2000). Mood, personality, and self-monitoring: negative affect and emotionality in relation to frontal lobe mechanisms of error monitoring. *Journal of Experimental Psychology: General, 129*(1), 43–60 Retrieved from. https://www.ncbi.nlm.nih.gov/pubmed/10756486.

Martiny-Huenger, T., Gollwitzer, P. M., & Oettingen, G. (2014). Distractor devaluation in a flanker task: Object-specific effects without distractor recognition memory. *Journal of Experimental Psychology: Human Perception and Performance, 40*(2), 613–625. http://dx.doi.org/10.1037/a0034130.

Moser, J. S., Moran, T. P., Schroder, H. S., Donnellan, M. B., & Yeung, N. (2013). On the relationship between anxiety and error monitoring: A meta-analysis and conceptual framework. *Frontiers in Human Neuroscience, 7*(August), 466. http://dx.doi.org/10.3389/fnhum.2013.00466.

Muth, C., & Carbon, C.-C. (2013). The aesthetic aha: On the pleasure of having insights into Gestalt. *Acta Psychologica, 144*(1), 25–30. http://dx.doi.org/10.1016/j.actpsy.2013.05.001.

Payne, K., & Lundberg, K. (2014). The affect misattribution procedure: Ten years of evidence on reliability, validity, and mechanisms. *Social and Personality Psychology Compass, 8*(12), 672–686. http://dx.doi.org/10.1111/spc3.12148.

Peirce, J. W. (2007). PsychoPy—Psychophysics software in python. *Journal of Neuroscience Methods, 162*(1–2), 8–13. http://dx.doi.org/10.1016/j.jneumeth.2006.11.017.

Peirce, J. W. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics, 2*. http://dx.doi.org/10.3389/neuro.11.010.2008.

Phillips, P. J., Moon, H., Rizvi, S. A., & Rauss, P. J. (2000). The FERET evaluation methodology for face-recognition algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 22*, 1090–1104. Retrieved from https://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=879790.

Phillips, P. J., Wechsler, H., Huang, J., & Rauss, P. J. (1998). The FERET database and evaluation procedure for face-recognition algorithms. *Image and Vision Computing, 16*(5), 295–306. Retrieved from https://www.sciencedirect.com/science/article/pii/S026288569700070X.

Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: is beauty in the perceiver's processing experience? *Personality and Social Psychology Review, 8*(4), 364–382. http://dx.doi.org/10.1207/s15327957pspr0804_3.

Satterthwaite, T. D., Ruparel, K., Loughead, J., Elliott, M. A., Gerraty, R. T., Calkins, M. E., ... Wolf, D. H. (2012). Being right is its own reward: Load and performance related ventral striatum activation to correct responses during a working memory task in youth. *NeuroImage, 61*(3), 723–729. http://dx.doi.org/10.1016/j.neuroimage.2012.03.060.

Schouppe, N., Braem, S., De Houwer, J., Silvetti, M., Verguts, T., Ridderinkhof, K. R., & Notebaert, W. (2014). No pain, no gain: The affective valence of congruency conditions changes following a successful response. *Cognitive, Affective, & Behavioral Neuroscience, 15*(1), 251–261. http://dx.doi.org/10.3758/s13415-014-0318-3.

Schroder, H. S., Moran, T. P., Infantolino, Z. P., & Moser, J. S. (2013). The relationship between depressive symptoms and error monitoring during response switching. *Cognitive, Affective, & Behavioral Neuroscience, 13*(4), 790–802. http://dx.doi.org/10.3758/s13415-013-0184-4.

Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology, 45*(3), 513–523.

Van de Cruys, S., & Wagemans, J. (2011). Putting reward in art: A tentative prediction error account of visual art. *I-Perception, 2*(9), 1035–1062. http://dx.doi.org/10.1068/i0466aap.

Yeung, N., Botvinick, M. M., & Cohen, J. D. (2004). The neural basis of error detection: Conflict monitoring and the error-related negativity. *Psychological Review, 111*(4), 931–959. http://dx.doi.org/10.1037/0033-295X.111.4.931.

Yeung, N., & Summerfield, C. (2012). Metacognition in human decision-making: confidence and error monitoring. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 367*(1594), 1310–1321. http://dx.doi.org/10.1098/rstb.2011.0416.

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist, 35*(2), 151–175. http://dx.doi.org/10.1037//0003-066X.35.2.151.

Zajonc, R. B. (2001). Mere exposure: A gateway to the subliminal. *Current Directions in Psychological Science, 10*(6), 224–228. http://dx.doi.org/10.1111/1467-8721.00154.