Inertia in value-driven attention

Ming-Ray Liao and Brian A. Anderson

Department of Psychological and Brain Sciences, Texas A&M University, College Station, Texas 77843, USA

Previously reward-associated stimuli persistently capture attention. We attempted to extinguish this attentional bias through a reversal learning procedure where the high-value color changed unexpectedly. Attentional priority shifted during training in favor of the currently high-value color, although a residual bias toward the original high-value color was still evident. Importantly, during a subsequent test phase, attention was initially more strongly biased toward the original high-value color, counter to the attentional priorities evident at the end of training. Our results show that value-based attentional biases do not quickly update with new learning and lag behind the reshaping of strategic attentional priorities by reward.

[Supplemental material is available for this article.]

The ability to selectively attend to important information in a complex visual scene, including reward-associated objects, is important for survival (Anderson 2013). The involuntary influence of reward history on attentional selection, referred to as value-driven attentional capture (VDAC), is well established (for reviews, see Anderson 2013, 2016a; Failing and Theeuwes 2018; Watson et al. 2019). Participants first learn to associate a feature like color (Anderson et al. 2011a,b), orientation (Theeuwes and Belopolsky 2012; Laurent et al. 2015), shape (Della Libera and Chelazzi 2009), or object category (Hickey et al. 2015; Donohue et al. 2016), with high value and another feature with low value. In a subsequent visual search task, previously reward-associated features capture attention even when explicitly task-irrelevant (e.g., Anderson et al. 2011a,b; Theeuwes and Belopolsky 2012).

VDAC has been shown to persist without additional reinforcement for well over one hundred trials (e.g., Anderson et al. 2011b, 2014; Anderson and Yantis 2012) and for up to 9 mo post-learning (Anderson and Yantis 2013), but eventually extinguishes with a sufficient number of nonreinforced trials (e.g., Anderson et al. 2011a, 2016). The persistence of VDAC is inconsistent with what might be predicted from classical conditioning, where the previously reward-predictive stimulus would cease to evoke a conditioned response more quickly with nonreinforcement (Pavlov 1927). Milner et al. (2020) found that VDAC is markedly slow to extinguish in a nonreinforced test phase, although it can eventually extinguish over many unrewarded trials, especially when the previously reward-associated feature appears more frequently. What remains to be understood are the dynamics by which value-based attentional priorities update when stimulus-reward contingencies change.

The present study investigated how a change in the relative value of different stimuli is reflected in the updating of attentional bias. We start with a training phase where participants learn to associate one color with high-value reward, one color with low-value reward, and another color with no reward. This is later followed by a second training phase that devalues the previously learned reward association: The former high-value color is now unrewarded and the previously unrewarded color is now predictive of high-value reward (reversal learning). Immediately following each of the two training phases is a test phase measuring attentional bias for the different colors. Humans can rapidly adapt to new reward

Corresponding author: m4liao@tamu.edu

Article is online at http://www.learnmem.org/cgi/doi/10.1101/lm.052027.120.

structures and even to frequently changing reward contingencies (e.g., Behrens et al. 2007; Ghahremani et al. 2010; Lee and Keramati 2017). We hypothesized that, although attentional capture by an originally high-value color may persist following such reversal learning, the bias should be markedly reduced, being overshadowed by the bias toward a newly learned high-value color (Fig. 1A). It is also possible that signal suppression occurs, whereby the original high-value color continues to generate a salience signal but is subsequently suppressed (Fig. 1B; Sawaki and Luck 2010; Gaspelin et al. 2015). Finally, to the degree that value-based attentional priority is computed over long periods of time and tracks total associated reward, the old and new high-value color would be expected to be similarly prioritized (Fig. 1C).

Forty participants were recruited from the Texas A&M University community. All reported normal or corrected-to-normal visual acuity, normal color vision, and provided written informed consent. Five participants did not complete the study (two because of experimenter error, two because of inability to reliably track the eyes, and one withdrew). The data for three participants were excluded from analyses due to low accuracy (either training or test phase <2.5 SD of the group mean). The final sample included 32 participants (17 females), with a mean age of 21.00 yr (SD=2.90), which indicated power (1– β)=0.82 using the oculomotor RT-based measures of VDAC in a prior eye tracking study (Anderson and Kim 2019b). Participants were compensated with their earnings from the task. All procedures were approved by the Texas A&M University Institutional Review Board and conformed with the principles outlined in the Declaration of Helsinki.

Stimuli were generated using MATLAB and Psychophysics Toolbox extensions (Brainard 1997), then presented on a Dell P2717H monitor linked to a Dell OptiPlex 7040. Participants viewed the monitor from a distance of \sim 70 cm in a dimly lit room. An EyeLink 1000 Plus desktop-mounted eye tracker (SR Research) monitored participants' right eye position.

Head position was maintained throughout the experiment using an adjustable chin and forehead rest (SR Research). Eye position was calibrated prior to each block of trials using nine-point calibration (Anderson and Yantis 2012; Liao and Anderson 2020) and

© 2020 Liao and Anderson This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first 12 months after the full-issue publication date (see http://learnmem.cshlp.org/site/misc/terms.xhtml). After 12 months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at http://creativecommons.org/licenses/by-nc/4.0/.

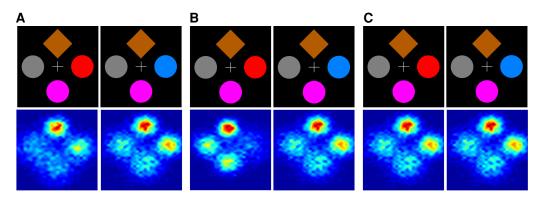


Figure 1. Schematic of experiment stimuli (*top*) and simulated priority map representation (*bottom*) for our hypotheses where old high-value color (red) loses its priority (*A*), the old high-value color is inhibited (*B*), and new (blue) and old high-value colors remain equally prioritized (*C*). The target is the shape singleton (diamond). More intense colors in the priority map reflect stronger attentional priority associated with the corresponding stimulus. The simulation is for illustrative purposes only, and the exact intensity values are arbitrary.

was manually drift-corrected by the experimenter as needed (each trial could only begin once a valid fixation had been registered within 1.2° of the center of the screen). During the presentation of the search array, the X and Y position of the eyes were continuously monitored in real time with respect to the six stimulus positions, such that fixations were coded online (Anderson and Kim 2019a,b).

Three colors were associated with monetary reward throughout various periods during the experiment. These critical colors were red (RGB: 255, 0, and 0), green (0, 255, and 0), and blue (0, 127, and 255), and served as the first or old high-value color, low-value color, and the second or new high-value color (randomly assigned for each participant). There were four other colors—gray (190, 190, and 190), pink (255, 0, and 255), yellow (240, 240, and 0), and brown (180, 90, and 0)—that served as neutral colors.

Each trial in the first training phase consisted of a fixation display, a search array, and a reward feedback display (Fig. 2A). The fixation cross $(1.1^{\circ}$ visual angle) remained on the screen until eye position had been registered within 1.2° of it for a continuous period of 500 msec. The search array consisted of four squares $(3.0^{\circ} \times 3.0^{\circ})$ on an imaginary circle with a radius of 8.3° . A region extending 1.1° beyond the boundary of each square was used to determine fixations. The reward feedback display consisted of the money earned on the current trial along with the updated total earnings.

Participants were instructed to fixate ("look directly at") either the red, green, or blue square on each trial and that different colors would be worth different amounts of money on average. One critical color (red, green, or blue) was associated with an 80% probability of a high reward of 8¢ and a 20% probability of a low reward of 2¢ (old high-value color), while for another color these percentages were reversed (low-value color), and the last color always resulted in 0¢. Only one of the critical colors was displayed on each trial, along with three neutral colors (randomly chosen without replacement). The position of each target color was counterbalanced across trials, the order of which was randomized. At no point in the experiment was any relationship between color and reward mentioned to participants. There were eight practice trials without a time limit, after which participants completed two blocks of 96 trials each (192 trials total).

Similar to the training phase, each trial in the first test phase consisted of a gaze-contingent fixation display and a search array in which the task was to fixate a target (Fig. 2B). The search array now consisted of either three circles $(1.7^{\circ} \text{ radius})$ and a diamond $(2.7^{\circ} \times 2.7^{\circ})$ or three diamonds and a circle, which were positioned on an imaginary circle with a radius of 9.8° . Regions extending 3.4° beyond the center of the circle and 3.0° beyond the center of the diamond were used as areas for determining stimulus fixations.

Participants were instructed to "pick the unique shape by looking directly at it," and to "try to be as fast as possible while still

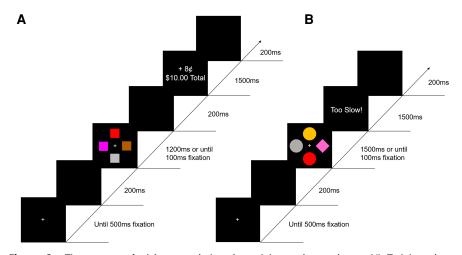


Figure 2. Time course of trial events during the training and test phases. (A) Training phase. Participants were tasked to fixate on the target color (red, green, or blue) on every trial. Participants had a 1200-msec time limit to fixate on the target for a continuous period of 100 msec before a blank was presented, followed by a feedback display showing the amount earned on the current trial along with total earnings. There was only one target color present on each trial; one was associated with a high-value reward when fixated, another with low-value reward, and the third with no reward. In the second training phase, the high-value color and the no reward color switched contingencies. (B) Test phase. Participants had 1500 msec to fixate on the unique shape (singleton) for a continuous period of 100 msec, while sometimes ignoring a critical color distractor from the training phase. If participants were successful in fixating the target within the time limit, the search array would be replaced with a blank screen for 500 msec, otherwise they would see the words "too slow!" for 1500 msec, followed by a 200-msec blank to end the trial. The second test phase was identical to the first.

being accurate." On neutral trials, the shapes appeared in the four neutral colors. On distractor-present trials, a nontarget shape would appear in one of the critical colors (red, green, or blue). The position of the distractor and the target shape was counterbalanced, while the neutral colors were selected randomly without replacement on each trial, and the order of trials was randomized. There was a 10-trial practice block with no time limit, after which there were two blocks of 96 trials (192 trials in total).

The second training phase was exactly the same as the first training phase, but the old high-value color was now never rewarded (always followed by $0\mathfrak{e}$) and the previously unrewarded color was now associated with an 80% probability of a high reward of $8\mathfrak{e}$ and a 20% probability of a low reward of $2\mathfrak{e}$ (new high-value color). There were also no more practice trials. The second test phase was exactly the same as the first test phase, with the exception that there were no more practice trials.

Response time (RT) was measured from the onset of the search array to the moment eye position entered into the fixation window surrounding the target, and only correct responses were included in the mean RT for each participant (82.5% and 90.3% of trials in training and test, respectively). RTs exceeding three SD of the mean for each condition for each participant were trimmed (1.75% of trials in the training phase and 1.5% in the test phase). RTs in the

test phase were normalized to the neutral condition (i.e., mean RT for the neutral condition was subtracted; see, e.g., Krebs et al. 2010; Liao et al. 2020) and RT for the critical colors in each phase were compared using three (target/distractor color) × 2 (block) repeated-measures analysis of variance (ANOVA).

In the first training phase, there was a significant main effect of target color, $F_{(2,62)}=6.98$, P=0.002, $\eta^2_P=0.184$, a significant main effect of block, $F_{(1,31)}=14.40$, P=0.001, $\eta^2_P=0.317$, and no interaction, $F_{(2,62)}=2.28$, P=0.111 (Fig. 3A). Post-hoc contrasts revealed that RT to the high-value target was significantly faster than to the other two targets (collapsed), $t_{(31)}=3.78$, P<0.001, d=0.67, which did not significantly differ from each other, t<1.

In the second training phase, there was again a significant main effect of target color, $F_{(2,62)} = 11.67$, P < 0.001, $\eta^2_P =$ 0.274, but no main effect of block, F < 1, and a marginally significant interaction, $F_{(2,62)} = 2.42$, P = 0.097, $\eta^2_P = 0.072$. Post-hoc contrasts revealed that, in the first block (block 3), RT was faster for both old high-value and new high-value targets compared with low-value targets, ts > 3.07, P < 0.005, ds > 0.54, and was similar for old high-value and new highvalue targets, t < 1. In the second block (block 4), new high-value targets were still reported faster than low-value targets, $t_{(31)} = 4.83$, P < 0.001, d = 0.85, whereas the difference between old high-value targets and low-value targets was now marginally significant, $t_{(31)} = 1.80$, P =0.081. Importantly, new high-value targets were now reported faster than old high-value targets, $t_{(31)} = 2.60$, P = 0.014, d = 0.46.

In the first test phase, there was no main effect of distractor condition, $F_{(2,62)} = 1.34$, P = 0.268, or a main effect of block or interaction, Fs < 1. Numerically, RT was slowest for the old high-value condition (Fig. 3B).

In the second test phase, the was no main effect of distractor condition, $F_{(2,62)}=1.98$, P=0.146, or block, F<1, but there was a significant interaction, $F_{(2,62)}=3.46$, P=0.038, $\eta^2_P=0.100$. Post-hoc contrasts revealed that in the first block (block 3), RT was slower in the old high-value distractor condition than in the other two conditions (collapsed), $t_{(31)}=2.75$, P=0.010, d=0.48, which did not significantly differ from each other, t<1. In the second block (block 4), there were no significant differences, although RT was now numerically slower in the new high-value condition. Distractor fixations were very infrequent (<3%) and there were no main effects or interactions in the frequency of distractor fixations in either test phase, P>0.08.

The findings of the present study extend our understanding of the persistence of VDAC. Our reversal learning manipulation was not immediately effective in modulating attentional bias in either the training or test phase. Following the change in value in the second training phase, the old high-value target continued to be prioritized during the initial block of trials (block 3), with a magnitude similar to that of the new high-value target. By the final block of

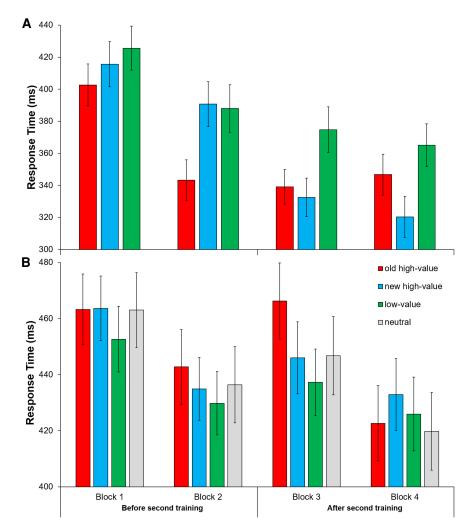


Figure 3. RT in the training phase for each target color (*A*) and in the test phase for each distractor color (*B*), for each block. Error bars represent standard error of the mean.

training, the new high-value target was significantly more prioritized, although the old high-value stimulus was still marginally more prioritized than the low-value target despite actually being worth less. Somewhat surprisingly, in spite of priority switching to the new high-value target, in the initial block of the final test phase (block 3), the old high-value rather than the new high-value distractor was the more prioritized stimulus, contrary to the hypotheses outlined in Figure 1. At least initially, VDAC did not mirror the priority given to task-relevant stimuli as a function of their value at the end of training, which would have predicted significantly stronger attentional capture by the new high-value stimulus. Our findings suggest that there is some inertia in the updating or recalibrating of value-based attentional priority.

Although participants successfully adapted to the new reward contingency, more quickly reporting the new high-value target by the end of training, this learning did not immediately generalize to the test phase. It is possible that attentional biases are not as flexible and malleable as learning and decision making, which have been shown to be sensitive to rapidly changing environments (Behrens et al. 2007; Ghahremani et al. 2010; Lee and Keramati 2017). The inflexibility of VDAC is what makes it a useful model for investigating clinical syndromes where valuable but undesirable stimuli conflict with behavioral goals (e.g., addiction and obesity) (Anderson 2016b). When an individual becomes addicted to a substance, related stimuli automatically capture attention similar to reward-associated distractors (for reviews, see Field and Cox 2008; Rooke et al. 2008), and continue to bias attention even within successfully recovered patients (Stormark et al. 1997; Marissen et al. 2006; Field and Cox 2008; Field et al. 2013). We see something similar in our results where the old high-value color continued to exert its influence on the attention system in spite of new learning.

It is possible that participants were learning to suppress inputs for the old high-value color in the second test phase and did not become proficient at doing so until later trials. Signal suppression is known to occur after capture under certain circumstances (Gaspelin et al. 2015; Gaspelin and Luck 2018; Sawaki and Luck 2010). This could explain the differences between the two blocks of the second test phase that may be before and after successful suppression of a color no longer associated with reward, which unfolds following exposure to the stimulus as an irrelevant distractor. It is also possible that automatic value-based attentional priority does not recalibrate unless such attending is explicitly counterproductive for some number of trials, as when previously reward-associated stimuli appear as task-irrelevant distractors (see Milner et al. 2020). More generally, our findings suggest that valuebased attentional biases are updated separately from the reshaping of reward-related attentional priorities evident in training phase performance, perhaps as a response to value-based distraction or continued nonreinforcement in the test phase, with old learning being capable of dominating over more recent learning even in the context of value-reversal.

Unexpectedly, we did not observe significant VDAC in the first test phase, although RT was numerically slower for the high-value than the low-value distractor. One notable difference between the present study and most prior studies of VDAC (e.g., Anderson et al. 2011a,b, 2014; Anderson and Kim 2019a,b) was that each training phase was shorter than the typical amount of training (192 trials divided across three color conditions versus at least 240 trials divided among only two color conditions). Since participants were not made aware of the color-reward contingencies, it might not be until the second training phase that the link between color and reward is fully grasped in the present study, with the violation of (possibly implicit) (see Grégoire and Anderson 2019) expectations from the first training phase serving to make the task contingencies more salient to participants. It is

also noteworthy that a sluggish bias for the old high-value color was evident during the second training phase, which is further consistent with robust learning of the original value associations exerting a persistent effect on attention.

There has been an increasing number of studies investigating the persistence of VDAC, but the present line of inquiry aims to provide insights into another way of overcoming attentional biases-with new learning. Our findings shed light on an important distinction between adjusting attentional priorities based on ongoing reward learning and the corresponding updating of involuntary attentional biases, with the former being more rapid and flexible, and the latter being slower and less malleable. Although it is unclear what test phase performance would have looked like if the second training phase had been longer, our results clearly show that value-based attentional priorities can shift during training before a corresponding bias emerges, with originally learned priorities significantly dominating attention when rewards are discontinued. In this respect, our findings offer compelling evidence that VDAC does not reflect a mere spillover of motivated attention from the training phase but rather involves a persistent and habitual bias (see Anderson 2016a) that cannot be quickly recalibrated with new learning.

References

- Anderson BA. 2013. A value-driven mechanism of attentional selection. *J Vis* 13: 7. doi:10.1167/13.3.7
- Anderson BA. 2016a. The attention habit: how reward learning shapes attentional selection. *Ann N Y Acad Sci* **1369:** 24–39. doi:10.1111/nyas .12957
- Anderson BA. 2016b. What is abnormal about addiction-related attentional biases? *Drug Alcohol Depend* **167:** 8–14. doi:10.1016/j.drugalcdep.2016 .08.002
- Anderson BA, Kim H. 2019a. On the relationship between value-driven and stimulus-driven attentional capture. Atten Percept Psychophys 81: 607– 613. doi:10.3758/s13414-019-01670-2
- Anderson BA, Kim H. 2019b. Test–retest reliability of value-driven attentional capture. *Behav Res Methods* **51:** 720–726. doi:10.3758/s13428-018-1079-7
- Anderson BA, Yantis S. 2012. Value-driven attentional and oculomotor capture during goal-directed, unconstrained viewing. Atten Percept Psychophys 74: 1644–1653. doi:10.3758/s13414-012-0348-2
- Anderson BA, Yantis S. 2013. Persistence of value-driven attentional capture. *J Exp Psychol* **39:** 6–9. doi:10.1037/a0030860
- Anderson BA, Laurent PA, Yantis S. 2011a. Learned value magnifies salience-based attentional capture. PLoS One 6: e27926. doi:10.1371/ journal.pone.0027926
- Anderson BA, Laurent PA, Yantis S. 2011b. Value-driven attentional capture. Proc Natl Acad Sci 108: 10367–10371. doi:10.1073/pnas.1104047108
- Anderson BA, Laurent PA, Yantis S. 2014. Value-driven attentional priority signals in human basal ganglia and visual cortex. *Brain Res* **1587**: 88–96. doi:10.1016/j.brainres.2014.08.062
- Anderson BA, Kuwabara H, Wong DF, Gean EG, Rahmim A, Braši JR, George N, Frolov B, Courtney SM, Yantis S. 2016. The role of dopamine in value-based attentional orienting. *Curr Biol* **26:** 550–555. doi:10.1016/j.cub.2015.12.062
- Behrens TE, Woolrich MW, Walton ME, Rushworth MF. 2007. Learning the value of information in an uncertain world. *Nat Neurosci* **10**: 1214–1221. doi:10.1038/nn1954
- Brainard DH. 1997. The psychophysics toolbox. *Spat Vis* **10:** 433–436. doi:10.1163/156856897X00357
- Della Libera C, Chelazzi L. 2009. Learning to attend and to ignore is a matter of gains and losses. *Psychol Sci* **20:** 778–784. doi:/10.1111/j.1467-9280 2009 02360 x
- Donohue SE, Hopf J-M, Bartsch MV, Schoenfeld MA, Heinze H-J, Woldorff MG. 2016. The rapid capture of attention by rewarded objects. *J Cogn Neurosci* **28:** 529–541. doi:10.1162/jocn_a_00917
- Failing M, Theeuwes J. 2018. Selection history: how reward modulates selectivity of visual attention. *Psychonom Bull Rev* **25:** 514–553. doi:10.3758/s13423-017-1380-y
- Field M, Cox WM. 2008. Attentional bias in addictive behaviors: a review of its development, causes, and consequences. *Drug Alcohol Depend* **97:** 1–20. doi:10.1016/j.drugalcdep.2008.03.030
- Field M, Mogg K, Mann B, Bennett GA, Bradley BP. 2013. Attentional biases in abstinent alcoholics and their association with craving. *Psychol Addict Behav* 27: 71–80. doi: 10.1037/a0029626

- Gaspelin N, Luck SJ. 2018. The role of inhibition in avoiding distraction by salient stimuli. *Trends Cogn Sci* **22:** 79–92. doi:10.1016/j.tics.2017.11
- Gaspelin N, Leonard CJ, Luck SJ. 2015. Direct evidence for active suppression of salient-but-irrelevant sensory inputs. Psychol Sci 26: 1740–1750. doi:10.1177/0956797615597913
- Ghahremani DG, Monterosso J, Jentsch JD, Bilder RM, Poldrack RA. 2010. Neural components underlying behavioral flexibility in human reversal learning. *Cereb Cortex* **20:** 1843–1852. doi:10.1093/cercor/bhp247
- Grégoire L, Anderson BA. 2019. Semantic generalization of value-based attentional priority. *Learn Mem* 26: 460–464. doi: 10.1101/lm.050336 .119
- Hickey C, Kaiser D, Peelen MV. 2015. Reward guides attention to object categories in real-world scenes. J Exp Psychol Gen 144: 264–273. doi:10 .1037/a0038627
- Krebs RM, Boehler CN, Woldorff MG. 2010. The influence of reward associations on conflict processing in the Stroop task. *Cognition* 117: 341–347. doi:10.1016/j.cognition.2010.08.018
- Laurent PA, Hall MG, Anderson BA, Yantis S. 2015. Valuable orientations capture attention. Vis Cogn 23: 133–146. doi:10.1080/13506285.2014 .965242
- Lee JJ, Keramati M. 2017. Flexibility to contingency changes distinguishes habitual and goal-directed strategies in humans. *PLoS Comput Biol* 13: e1005753. doi:10.1371/journal.pcbi.1005753
- Liao MR, Anderson BA. 2020. Reward learning biases the direction of saccades. *Cognition* **196**: 104145. doi:10.1016/j.cognition.2019.104145 Liao MR, Grégoire L, Anderson BA. 2020. The influence of threat and
- Liao MR, Grégoire L, Anderson BA. 2020. The influence of threat and aversive motivation on conflict processing in the Stroop task. Atten Percept Psychophys 82: 2802–2813. doi:10.3758/s13414-020-02072-5
- Marissen MAE, Franken IHA, Waters AJ, Blanken P, van den Brink W, Hendriks VM. 2006. Attentional bias predicts heroin relapse following

- treatment. *Addiction* **101:** 1306–1312. doi: 10.1111/j.1360-0443.2006 .01498.x
- Milner AE, MacLean MH, Giesbrecht B. 2020. The persistence of value-driven attention capture is task dependent. PsyArxiv doi:10 .31234/OSF.IO/967K8
- Pavlov I. 1927. *Conditioned reflexes (translated by GV Anrep)*. Oxford University Press, Oxford, London.
- Rooke SE, Hine DW, Thorsteinsson EB. 2008. Implicit cognition and substance use: a meta-analysis. *Addict Behav* 33: 1314–1328. doi:10 .1016/j.addbeh.2008.06.009
- Sawaki R, Luck SJ. 2010. Capture versus suppression of attention by salient singletons: electrophysiological evidence for an automatic attend-to-me signal. Atten Percept Psychophys 72: 1455–1470. doi:10.3758/app.72.6 .1455
- Stormark KM, Field NP, Hugdahl K, Horowitz M. 1997. Selective processing of visual alcohol cues in abstinent alcoholics: an approach-avoidance conflict? *Addict Behav* 22: 509–519. doi:10.1016/s0306-4603(96) 00051-2
- Theeuwes J, Belopolsky AV. 2012. Reward grabs the eye: oculomotor capture by rewarding stimuli. *Vis Res* **74:** 80–85. doi:10.1016/j.visres.2012.07
- Watson P, Pearson D, Wiers RW, Le Pelley ME. 2019. Prioritizing pleasure and pain: attentional capture by reward-related and punishment-related stimuli. *Curr Opin Behav Sci* **26:** 107–113. doi:10.1016/j.cobeha.2018.12 .002

Received May 28, 2020; accepted in revised form September 1, 2020.