Semantic generalization of value-based attentional priority
Laurent Grégoire and Brian A. Anderson
Department of Psychology, Texas A&M University College Station, Texas 77843-4235, USA

This study aimed to determine whether attentional prioritization of stimuli associated with reward transfers across conceptual knowledge independently of physical features. Participants successively performed two color-word Stroop tasks. In the learning phase, neutral words were associated with high, low, or no monetary reward. In the generalization phase (in which no reward was delivered), synonyms of words previously paired with reward served as Stroop stimuli. Results are consistent with semantic generalization of stimulus-reward associations, with synonyms of high-value words impairing color-naming performance, although this effect was particular to participants who were unaware of the reward contingencies.

Attention prioritizes stimuli previously paired with reward independently of current goals and perceptual salience (e.g., Anderson et al. 2011, 2014a; Anderson and Halpern 2017; for review, see Watson et al. 2019). This value-based attentional priority can transfer to novel stimuli that share a defining feature (e.g., color) with previously rewarded items (Anderson et al. 2012; see also Mine and Saiki 2015, 2018). Hickey and Peelen (2015, 2017) also reported a generalization of associations between naturalistic visual stimuli and reward to objects of the same category (e.g., cars). More recently, Andreatta and Pauli (2019) showed that appetitive conditioned responses transferred to stimuli perceptually related to signals paired with food consumption in a prior acquisition phase. However, studies about the generalization of stimulus-reward associations mainly focused on perceptual cues. The potential semantic generalization of such associations is unknown, even though real-world learning situations, especially emotional experiences, often entail conceptual knowledge (Dunsmoor and Murphy 2015).

Using the color-word Stroop paradigm, Krebs et al. (2010) observed an enhanced interference effect when the semantic meaning of words printed in an incongruent color (e.g., the word "green" printed in red) corresponded to the ink color (e.g., green) that signaled reward availability. This result suggests that the color-reward association transferred to the color word, but if so, it is not a semantic generalization per se. At most, the study of Krebs et al. evidences the transfer of an association between a perceptual representation (i.e., the color) and reward toward a semantic representation (i.e., the color-word meaning). A semantic generalization would be demonstrated by the transfer of an association between a semantic representation and reward toward another (different but conceptually related) semantic representation (Razran 1939; Staats et al. 1959; Paciorek and Williams 2015).

Semantic generalization was reported in the context of fear conditioning using electrical stimulation (Dunsmoor et al. 2012; Dunsmoor and Murphy 2014; Boyle et al. 2016; Grégoire and Greening 2019). However, semantic generalization has apparently never been investigated in the domain of attention and no study focused on value-based effects. Whether stimulus-reward associations, and their effect on attention, generalize to semantically related stimuli thus remains an open question. This potential process could provide valuable insight into a critical aspect of adaptation (i.e., detect stimuli associated with reward) and improve the understanding and treatment of maladaptive behaviors to which attentional biases contribute (e.g., substance abuse; Field and Cox 2008; Anderson 2016b).

The present study aimed to determine whether attentional prioritization of stimuli associated with reward transfers across conceptual knowledge independently of perceptual features. We devised a color-word Stroop task in which neutral words were paired with high, low, or no monetary reward during a learning phase. In a subsequent generalization phase, participants performed a similar task with synonyms of words previously paired with reward. We hypothesized that synonyms of words paired with high reward would produce a Stroop interference effect (i.e., would slow down the color-identifying task), relative to synonyms of words paired with low or no reward, because they should be prioritized by attention (due to their semantic association with words related to high reward) and more difficult to inhibit. Prior research suggests that semantic generalization of fear conditioning may be particular to participants aware of the stimulus-outcome contingencies (Grégoire and Greening 2019), while another study found that reward-related interference was particular to unaware participants (Leganes-Fonteneau et al. 2019). A secondary research question therefore aimed to assess whether awareness of stimulus-reward associations modulated the potential semantic generalization of value-based attentional priority.

Thirty-six participants were recruited from the Texas A&M University community. All were native English speakers, reported normal or corrected-to-normal visual acuity and normal color vision. Data from one participant was discarded because he reported using strategies to avoid reading words (by squinting). Two additional participants were removed from analyses due to a low proportion of correct responses (below 2.5 SD of the group mean). The final sample included 33 participants (21 female, mean age = 22.03 yr [SD = 3.60]). All procedures were approved by the Texas A&M University Institutional Review Board and were conducted

Corresponding author: lgregoire1@exchange.tamu.edu
Article is online at http://www.learnmem.org/cgi/doi/10.1101/lm.050336.119.

© 2019 Grégoire 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/.

26/460–464; Published by Cold Spring Harbor Laboratory Press
ISSN 1549-5485/19; www.learnmem.org Learning & Memory
in accordance with the principles expressed in the Declaration of Helsinki.

Six pairs of synonyms were selected from The University of South Florida Word Association, Rhyme and Word Fragmentation Norms database of free association (Nelson et al. 1998): tuna-fish, clock-time, oak-tree, assist-help, fuel-gas, yolk-egg. The chosen pairs were all rated highly (i.e., above 65%) for the frequency of free association when single word priming was provided. There was no phonological or orthographic similarity between either word of each pair. The first word in each pair was presented in the learning phase, and their corresponding synonyms in the generalization phase. The learning phase consisted of three conditions: high reward, low reward, and no reward. Each condition comprised two words (which were always paired together in a specific condition to prevent grouping by category, e.g., tuna and yolk in the category food): tuna and clock, oak and assist, or fuel and yolk. These three pairs of words and conditions were fully crossed and counterbalanced across participants. For the sake of simplicity, we kept the same terminology for the three conditions (high reward, low reward, and no reward) in the two phases, though participants did not receive reward in the generalization phase. Words were presented in equiluminant red, green, blue, and purple.

Before the learning phase, participants performed 24 practice trials with six neutral words (different and not semantically related to the experimental words) presented in each of the four colors. The learning phase and the generalization phase were split into three and two 120-trial blocks, respectively. Each block comprised an equal number of high-, low-, and no-reward trials (i.e., 40). In each block, each of the six words was presented five times in each of the four colors (Fig. 1); trials were pseudorandomly ordered, excluding immediate repetitions of colors and words.

Participants were instructed to report the ink color of each word as quickly and accurately as possible, ignoring their meaning, by pressing the “C” key if the word was colored in green or purple or the “M” key if the word was colored in blue or red. Before the learning phase, we specified that correct responses sometimes resulted in the delivery of a monetary reward (no information about stimulus–reward contingencies was given). Before the generalization test, participants were informed that no reward was delivered in this phase. However, to maintain motivation, we indicated that they would receive a $3 bonus if their accuracy was >90%. Upon completion of the experiment, participants were given the cumulative monetary reward they had earned.

After the generalization phase, participants provided self-reported evaluations of their contingency awareness between words presented in the learning phase and reward. Each of the six words was presented once in each of the four colors, leading to 24 trials. Stimuli were pseudorandomly ordered and displayed around the center of the screen in the same way as in the learning phase. At each trial, participants were asked: “How much money do you think you would make for a correct response to this item?” and selected 10¢, 2¢, or 0¢ (three-alternative forced-choice) by clicking on the amount with the computer mouse. Contingency-awareness measures aimed to determine both if participants were aware of the stimulus–reward contingencies and if awareness was correlated with Stroop effects. At the end of the experiment, participants responded to a short questionnaire to indicate if they used strategies during the Stroop task.

Misses and errors represented, respectively, 0.78% and 3.93% of the trials in the learning phase, and 0.28% and 3.50% of the trials in the generalization phase (see Table 1). Response times (RTs) for correct responses beyond three standard deviations of the mean for each participant (1.27% in the learning phase and 0.69% in the generalization phase) were trimmed (Grégoire et al. 2013, 2014, 2015).

A 3 × 3 repeated-measures analyses of variance (ANOVA) conducted on the proportion of errors and misses measured in the learning phase with condition (high reward, low reward, no reward) and block (1, 2, 3) as within-subject variables revealed no main effects or interactions, Fs < 1.9, Ps > 0.13. The same ANOVA performed on correct RTs revealed no main effects or interaction, Fs < 1.8, Ps > 0.17.

A 3 × 2 repeated-measures ANOVA conducted on the proportion of errors and misses measured in the generalization phase with condition (high reward, low reward, no reward) and block (1, 2) as within-subject variables revealed a marginal main effect of condition, $F_{(2,32)} = 3.03, P = 0.056, \eta^2_p = 0.086$, a significant main effect of block in which errors and misses decreased with experience, $F_{(2,32)} = 9.69, P = 0.004, \eta^2_p = 0.232$, and no interaction, $F_{(2,64)} = 0.88, P = 0.420$. Subsequent t-tests indicated that, compared to the no-reward condition, the proportion of errors and misses was significantly lower in the high-reward condition, $t_{(32)} = 2.16, P = 0.039, d = 0.38$, and marginally lower in the low-reward condition, $t_{(32)} = 2.02, P = 0.052, d = 0.35$ (no significant effect was observed between high- and low-reward conditions, t < 1). The same ANOVA performed on correct RTs revealed no main effects or interaction, Fs < 1.9, Ps > 0.16.

From the contingency-awareness measures, we computed difference scores
in the high-reward condition, faster in the no-reward condition than the unaware group, RTs were significantly (or marginally) greater in the unaware group than in the aware group, high- vs. no-reward: $t_{(31)} = 5.34, P < 0.001, d = 1.88$, high- vs. low-reward: $t_{(31)} = 2.02, P = 0.052, d = 0.65$, low- vs. no-reward: $t_{(31)} = 2.79, P = 0.009, d = 1.03$.

Similar analyses performed on errors and misses in the learning phase revealed that the Stroop effect for the difference between high- and low-reward conditions was marginally greater in the aware group than in the unaware group, $t_{(31)} = 2.03, P = 0.051, d = 0.672$. All the remaining effects were nonsignificant (all $P > 0.10$).

The current study aimed to determine whether the effects of reward learning on attentional priority generalize to semantically related but perceptually unrelated stimuli. The striking result is a negative correlation between the Stroop effect measured on RTs in the generalization phase and contingency awareness. Thus, awareness of word-reward contingencies presented during learning (for high- and no-reward conditions) predicted the amplitude of the Stroop effect (RT high reward–RT no reward) in the generalization phase.

Table 1. Correct response times and proportion of errors and misses as a function of experimental conditions in learning and generalization phases

|                                      | High reward | Low reward | No reward |
|--------------------------------------|-------------|------------|-----------|
| **Learning phase**                  |             |            |           |
| Proportion of errors and misses      | 4.87 (4.76) | 4.55 (3.86) | 4.72 (3.85) |
| Correct response times (msec)        | 524.23 (63.37) | 527.28 (65.93) | 524.18 (67.49) |
| **Generalization phase**             |             |            |           |
| Proportion of errors and misses      | 3.41 (2.45) | 3.41 (2.24) | 4.51 (3.12) |
| Correct response times (msec)        | 504.58 (58.61) | 504.17 (57.37) | 504.68 (60.91) |

Note. Standard deviations are in parentheses.

Unsurprisingly, two-by-two comparisons for contingency-awareness scores of the three experimental conditions (i.e., high reward–no reward, high reward–low reward, and low reward–no reward) revealed nonsignificant effects in the unaware group (all $P > 0.10$) and significant effects in the aware group (all $P < 0.001$).

To clarify the negative correlations observed between Stroop effects and contingency-awareness effects in the generalization phase, we analyzed Stroop effects separately for each group. In the unaware group, RTs were significantly faster in the no-reward condition than in the high-reward condition, $t_{(19)} = 3.48, P = 0.003, d = 0.78$. No significant effect was observed between the low-reward condition and the other two conditions, $t_s < 1.4, P_s > 0.18$ (Fig. 2B). In the aware group, RTs were significantly slower in the no-reward condition than in the low- and high-reward conditions, $t_{(12)} = 2.90, P = 0.013, d = 0.80$, and $t_{(12)} = 3.91, P = 0.002, d = 1.08$, respectively. RTs were also marginally slower in the low-reward condition than in the high-reward condition, $t_{(12)} = 1.80, P = 0.097, d = 0.50$ (Fig. 2C). Between-group comparisons revealed that Stroop effects were significantly (or marginally) greater in the unaware group than in the aware group, high- vs. no-reward: $t_{(31)} = 5.34, P < 0.001, d = 1.88$, high- vs. low-reward: $t_{(31)} = 2.02, P = 0.052, d = 0.65$, low- vs. no-reward: $t_{(31)} = 2.79, P = 0.009, d = 1.03$.

The striking result is a negative correlation between the Stroop effect measured on RTs in the generalization phase and contingency awareness. Thus, awareness of word-reward contingencies presented during learning (for high- and no-reward conditions) predicted the amplitude of the Stroop effect (RT high reward–RT no reward) in the generalization phase.

Figure 2. Generalization phase. (A) Relationship between Stroop effect (RT high reward–RT no reward) and contingency-awareness effect (high reward score–no reward score). Correct response times as a function of condition (high reward, low reward, no reward) in (B) the unaware group and (C) the aware group. Asterisks indicate significant differences (*$P < 0.05$, **$P < 0.01$), and the dagger indicates a marginally significant difference (†$P < 0.10$). Error bars depict within-subjects confidence intervals calculated using the Cousineau method (Cousineau 2005) with a Morey correction (Morey 2008).
When analyzed separately for each group, results revealed two opposite patterns for unaware and aware participants in the generalization phase. We observed a pattern consistent with value-driven interference in the unaware group, with slower RTs in the high-reward condition than in the no-reward condition. Thus, generalized words captured more attention when associated with high reward rather than no reward in participants unaware of the relationship between conditioned words and reward. This finding, which may seem counterintuitive, fits nicely with prior studies of value-driven attention in which participants as a group are largely unaware of the reward contingencies (Theeuwes and Belopolsky 2012; Anderson 2015), and a recent study in which reward-related interference was particular to unaware participants (Leganes-Fonteneau et al. 2019). Our data also showed an unexpected negative Stroop effect in the aware group; although these results should be interpreted with caution, given the small sample size (N = 13), they potentially reflect value-based signal suppression (Gaspenl and Luck 2018). According to the signal suppression hypothesis (Sawaki and Luck 2010), a top-down control mechanism may prevent attentional capture and reduce the processing of salient stimuli below baseline levels. Hickey et al. (2010) suggested that the salience of a stimulus increases after pairing with reward (Anderson and Kim 2019). Thus, generalized words associated with high reward could be perceptually more salient than generalized words associated with low and no reward (due to transfer of word-reward associations acquired during learning). In order to perform the task more efficiently, aware participants might have actively suppressed the more perceptually salient reward-associated words.

It is worth noting that the assessment of contingency-awareness was performed after the Stroop task, whereas concurrent measures, taken during the experimental phase, are usually more sensitive (Lovibond and Shanks 2002). Our approach therefore provides a conservative measure of awareness. We opted for post-experimental measures to prevent or reduce the possibility that participants find out the goal of our study. A trial-by-trial assessment of awareness might have informed participants about our objectives and biased our results. For instance, this could have tempted participants to discover relationships between stimuli and reward resulting in an artificially high proportion of aware participants.

In the learning phase, errors and misses tended to be more frequent in the high-reward condition for aware participants, relative to the low-reward condition, suggesting that words associated with high reward were more distracting in the aware group, potentially due to sign-tracking (Le Pelley et al. 2015). However, RTs in the learning phase did not differ between conditions (even when analyzed separately for each group). That effects of reward on distraction were more robust in a test phase following learning is a common finding in the literature (Anderson et al. 2012; Mine and Saiki 2015; Anderson 2016a; Anderson and Halpern 2017), and it is possible that participants only learned the contingencies toward the end of training (see Anderson et al. 2014b).

To conclude, our results are consistent with a semantic generalization of stimulus–reward associations independent of perceptual features; however, effects observed with generalized stimuli depend on the awareness of the relationship between conditioned stimuli and reward.

References
Anderson BA. 2015. Value-driven attentional priority is context specific. Psychon Bull Rev 22: 750–756. doi:10.3758/s13423-014-0724-0
Anderson BA. 2016a. Social reward shapes attentional biases. Cogn Neurosci 7: 30–36. doi:10.1080/17588928.2015.1047823
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, Halpern M. 2017. On the value-dependence of value-driven attentional capture. Atten Percep Psychophys 79: 1001–1011. doi:10.3758/s13414-017-1289-6
Anderson BA, Kim H. 2019. On the relationship between value-driven and stimulus-driven attentional capture. Atten Percep Psychophys 81: 607–613. doi:10.3758/s13414-019-01670-2
Anderson BA, Laurent PA, Yantis S. 2011. Value-driven attentional capture. Proc Natl Acad Sci 108: 10367–10371. doi:10.1073/pnas.1104074108
Anderson BA, Laurent PA, Yantis S. 2012. Generalization of value-based attentional priority. Vis cogn 20: 647–658. doi:10.1080/13506285.2012.679711
Anderson BA, Laurent PA, Yantis S. 2014a. Value-driven attentional priority signals in human basal ganglia and visual cortex. Brain Res 1587: 88–96. doi:10.1016/j.brainres.2014.06.082
Anderson BA, Leal SL, Hall MG, Yassa MA, Yantis S. 2014b. The attribution of value-based attentional priority in individuals with depressive symptoms. Cogn Affect Behan Neurosci 14: 1221–1227. doi:10.3758/s13415-014-0303-2
Andreatta M, Pauli P. 2019. Generalization of appetitive conditioned stimuli. Psychophysiology 56: e13397. doi:10.1111/psyp.13397
Ben-Haim MS, Mama Y, Icht M, Algom D. 2014. Is the emotional Stroop task a special case of mood induction? Evidence from sustained effects of attention under emotion. Atten Percep Psychophys 76: 81–97. doi:10.3758/s13414-013-0545-7
Boyle S, Roche B, Dymond S, Hermans D. 2016. Generalisation of fear and avoidance along a semantic continuum. Cogn Emot 30: 340–352. doi:10.1080/02699931.2014.1085811
Cousineau D. 2005. Confidence intervals in within-subject designs: a simpler solution to Loftus and Masson’s method. Tutor Quant Methods Psychol 1: 42–45. doi:10.20982/tqmp.01.1.0042
Dunsmoor JE, Murphy GL. 2014. Stimulus typicality determines how broadly fear is generalized. Psychol Sci 25: 1816–1821. doi:10.1177/0956797614535401
Dunsmoor JE, Murphy GL. 2015. Categories, concepts, and conditioning: how humans generalize fear. Trends Cogn Sci 19: 73–77. doi:10.1016/j.tics.2014.12.003
Dunsmoor JE, Martin A, Labar KS. 2012. Role of conceptual knowledge in learning and retention of conditioned fear. Biol Psychol 89: 300–305. doi:10.1016/j.biopsycho.2012.06.009
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
Gaspenl 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.2017.11.001
Grégoire L, Greening SG. 2019. Fear of the known: semantic generalisation of fear conditioning across languages in bilinguals. Cogn Emot 35: 1324–1338. doi:10.1080/02699931.2019.1604319
Grégoire L, Perruchet P, Poulin-Charronnat R. 2013. The musical Stroop effect opening a new avenue to research on automatisms. Exp Psychol 60: 269–278. doi:10.1027/1618-3169/a000197
Grégoire L, Perruchet P, Poulin-Charronnat R. 2014. About the unidirectionality of interference: insight from the musical Stroop effect. Q J Exp Psychol 67: 2071–2089. doi:10.1080/17470218.2014.889392
Grégoire L, Perruchet P, Poulin-Charronnat B. 2015. How does Stroop interference change with practice? A reappraisal from the musical Stroop paradigm. J Exp Psychol Learn Mem Cogn 41: 417–425. doi:10.1037/xlm0000044
Hickey C, Peelen MV. 2015. Neural mechanisms of incentive salience in naturalistic human vision. Neuron 85: 512–518. doi:10.1016/j.neuron.2014.12.049
Hickey C, Peelen XV. 2017. Reward selectively modulates the lingering neural representation of recently attended objects in natural scenes. J Neurosci 37: 7297–7304. doi:10.1523/JNEUROSCI.0684-17.2017
Hickey C, Chelazzi L, Theeuwes J. 2010. Reward changes salience in human vision via the anterior cingulate. J Neurosci 30: 11096–11103. doi:10.1523/JNEUROSCI.0656-10.2010
Krebs RM, Boehler CN, Woldorf 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
Leganes-Fonteneau M, Nikolau K, Scott R, Duka T. 2019. Knowledge about the predictive value of reward conditioned stimuli modulates their interference with cognitive processes. Learn Mem 26: 66–76. doi:10.1101/lm.048272.118
Le Pelley ME, Pearson D, Griffiths O, Beesley T. 2015. When goals conflict with values: counterproductive attentional and oculomotor capture by reward-related stimuli. J Exp Psychol Gen 144: 158–171. doi:10.1037/xge0000037
Lovibond PE, Shanks DR. 2002. The role of awareness in Pavlovian conditioning: empirical evidence and theoretical implications. *J Exp Psychol Anim Behav Process* **28**: 3–26. doi:10.1037//0097-7403.28.1.3

Mine C, Saiki J. 2015. Task-irrelevant stimulus-reward association induces value-driven attentional capture. *Atten Percept Psychophys* **77**: 1896–1907. doi:10.3758/s13414-015-0894-5

Mine C, Saiki J. 2018. Pavlovian reward learning elicits attentional capture by reward-associated stimuli. *Atten Percept Psychophys* **80**: 1083–1095. doi:10.3758/s13414-018-1502-2

Morey RD. 2008. Confidence intervals from normalized data: a correction to Cousineau (2005). *Tutor Quant Methods Psychol* **4**: 61–64. doi:10.20982/tqmp.04.2.p061

Nelson DL, McEvoy CI, Schreiber TA. 1998. *The University of South Florida word association, rhyme, and word fragment norms*. Retrieved from http://w3.usf.edu/FreeAssociation

Pactorek A, Williams JN. 2015. Semantic generalization in implicit language learning. *J Exp Psychol Learn Mem Cogn* **41**: 989–1002. doi:10.1037/xlm0000100

Razran GHS. 1939. A quantitative study of meaning by a conditioned salivary technique (Semantic conditioning). *Science* **90**: 89–90. doi:10.1126/science.90.2326.89-a

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**: 1453–1470. doi:10.3758/app.72.6.1455

Staats AW, Staats CK, Heard WG. 1939. Language conditioning of meaning to meaning using a semantic generalization paradigm. *J Exp Psychol* **57**: 187–192. doi:10.1037/h0042274

Theeuwes J, Belopolsky AV. 2012. Reward grabs the eye: oculomotor capture by rewarding stimuli. *Vis Res* **74**: 80–85. doi:10.1016/j.visres.2012.07.024

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 July 22, 2019; accepted in revised form October 1, 2019.