CrossMark

Value-driven attentional capture in the auditory domain

Brian A. Anderson[1]

Abstract It is now well established that the visual attention system is shaped by reward learning. When visual features are associated with a reward outcome, they acquire high priority and can automatically capture visual attention. To date, evidence for value-driven attentional capture has been limited entirely to the visual system. In the present study, I demonstrate that previously reward-associated sounds also capture attention, interfering more strongly with the performance of a visual task. This finding suggests that value-driven attention reflects a broad principle of information processing that can be extended to other sensory modalities and that value-driven attention can bias cross-modal stimulus competition.

Key words Selective attention · Auditory attention · Cross-modal attention · Reward learning

Attention determines which among multiple competing stimuli are represented in the brain (Desimone & Duncan, 1995). Attentional selection has long been understood to arise from the interplay between the bottom-up physical salience of the stimulus (e.g., Theeuwes, 1992; Yantis & Jonides, 1984) and its relationship to the top-down goals of the observer (e.g., Folk, Remington, & Johnston, 1992). More recently, a wealth of research has demonstrated that reward information provides a third source of input to the attention system. When a large reward is received, attention is strongly primed to select the rewarded target (Hickey, Chelazzi, & Theeuwes, 2010; see also Della Libera & Chelazzi, 2006). More persistent

✉ Brian A. Anderson
bander33@jhu.edu

[1] Psychological & Brain Sciences, Johns Hopkins University, 3400 N. Charles St., Baltimore, MD 21218-2686, USA

attentional biases develop for stimuli that have been learned to predict a reward outcome (e.g., Anderson, Laurent, & Yantis, 2011a, b; Della Libera & Chelazzi, 2009; Raymond & O'Brien, 2009). Such previously reward-associated stimuli capture attention even when task-irrelevant and physically nonsalient (e.g., Anderson et al., 2011b), demonstrating that learned value plays a distinct role in the guidance of attention (referred to as *value-driven attention;* see Anderson, 2013, for a review).

Value-driven attention was originally identified in the domain of vision, and a large effort has been undertaken to characterize the influence of learned value on visual attention. This includes its spatial specificity (e.g., Anderson & Yantis, 2012; Failing & Theeuwes, 2014; Theeuwes & Belopolsky, 2012), contextual specificity (Anderson, 2015a, b; Anderson, Laurent, & Yantis, 2012), extension to different visual properties (e.g., Chelazzi et al., 2014; Laurent, Hall, Anderson, & Yantis, 2015; Lee & Shomstein, 2014), persistence (Anderson & Yantis, 2013), relationship with psychopathology (Anderson, Faulkner, Rilee, Yantis, & Marvel, 2013; Anderson, Leal, Hall, Yassa, & Yantis, 2014), mechanism of learning (e.g., Sali, Anderson, & Yantis, 2014; Le Pelley, Pearson, Griffiths, & Beesley, 2015), and neural mechanisms (e.g., Anderson, Laurent, & Yantis, 2014; Hickey & Peelen, 2015; Peck, Jangraw, Suzuki, Efem, & Gottlieb, 2009; Qi, Zeng, Ding, & Li, 2013).

Principles of salience-driven and goal-directed attention, although most extensively studied in vision, have been extended to other sensory modalities (see Spence, 2010, for a review). In the auditory domain, for example, a task-irrelevant sound can draw spatial attention to its approximate source of origin (e.g., Spence & Driver, 1994, 1997). Salient auditory singletons can capture attention and interfere with the ability to discriminate an auditory target (Dalton & Lavie, 2004, 2007), similar to the impact of salient singletons in the visual

domain (e.g., Bacon & Egeth, 1994; Folk & Anderson, 2010; Theeuwes, 1992). Auditory singletons can also capture attention and interfere with the performance of a nonspatial visual search task when they share a defining feature with the visual target (e.g., both are singletons in their duration; Dalton & Spence, 2007). In addition, identifying a target sound can elicit an attentional blink that interferes with subsequent visual processing (Arnell & Jolicoeur, 1999), although the robustness of such temporal cross-modal interference may be limited (e.g., Duncan, Martins, & Ward, 1997). Shifting attention from processing auditory information to processing visual information and vice versa incurs a performance cost and recruits the same parietal mechanisms known to mediate attention shifts within a sensory modality (Shomstein & Yantis, 2004; Yantis et al., 2002).

Learned associations with reward can modulate the sensory processing of a stimulus within modalities other than vision (e.g., Pantoja et al., 2007), and previously reward-associated sounds can facilitate the processing of visual stimuli through cross-modal interactions (Pooresmaeili et al., 2014). There is also evidence that auditory stimuli can influence value-driven attention in the visual domain. Specifically, when a sound is used to indicate reward, visual stimuli that are paired with this sound subsequently capture attention (Miranda & Palmer, 2014). However, it is currently unknown whether the principle of value-driven attention similarly extends to sensory modalities beyond vision, such that previously reward-associated sounds automatically capture attention. Relatedly, it remains unknown whether value-driven attention can bias cross-modal stimulus competition, with previously reward-associated sounds interfering with the performance of a visual task. Evidence in the affirmative would suggest that value-driven attention reflects a broad principle of human information processing.

In the present study, I examine whether previously reward-associated sounds can automatically capture attention and interfere with the performance of a visual task. Across two experiments, I demonstrate that a sound previously associated with high reward interferes more strongly with the identification of a visual target than a sound previously associated with comparatively low reward. My findings extend the principle of value-driven attentional capture into the domain of auditory attention.

Experiment 1

In Experiment 1, participants first completed a training phase involving the detection of auditory targets. Participants listened for one of two spoken letters played over the computer speakers and pressed a key every time they heard one of those target sounds. One target sound yielded a high reward every time it was identified, whereas the other yielded a low reward.

Nontarget spoken letters were also played, to which participants withheld responding. Irrelevant visual letters were presented during training to compete with the perception of the spoken letters (see Desimone & Duncan, 1995), requiring participants to select auditory stimuli while explicitly ignoring concurrent visual stimulation. This was to maximize the strength of any learned attentional bias on subsequent information processing, which could involve both enhanced activation of the target sound as well as the suppression of concurrent visual input.

Once participants had experienced these sound-reward associations during training, these same sounds served as task-irrelevant distractors during a visual search task for a shape-defined target in a subsequent unrewarded test phase. The distractor sound could be that of a former high-reward target (high-value distractor), a former low-reward target (low-value distractor), or a former nontarget. All visual search trials involved a distractor sound in order to equate the alerting effects that auditory stimuli might have on task performance across conditions. Of interest was whether the high-value distractor sound impaired performance of the visual search task, consistent with cross-modal value-driven attentional capture in the auditory domain.

Method

Participants Twenty-six participants were recruited from the Johns Hopkins University community. All reported normal or corrected-to-normal visual acuity and normal color vision. One participant was an outlier, with a capture score that deviated from the grand mean by more than 2.5 standard deviations and was replaced; replacing this participant did not change any of the statistical conclusions.

Apparatus A MacBook Pro laptop computer equipped with MATLAB software and Psychophysics Toolbox extensions (Brainard, 1997) was used to present the stimuli. Participants were seated approximately 65 cm from the laptop in a dimly lit room. Manual responses were entered using the keyboard on the laptop. The sounds were presented from the built-in speakers.

Training phase

Stimuli A steady stream of colored letters (each 1.3° × 1.5° visual angle) was presented at the center of the screen (see Fig. 1a). The identity of each letter was selected from the set {A, C, F, G, H, J, K, M, N, P, R, T, U, V, X, Y}, and the color of each letter was selected from the set {blue, cyan, pink, orange, yellow, white}. On some letter displays, simultaneous with the onset of the visual letter, a spoken letter was
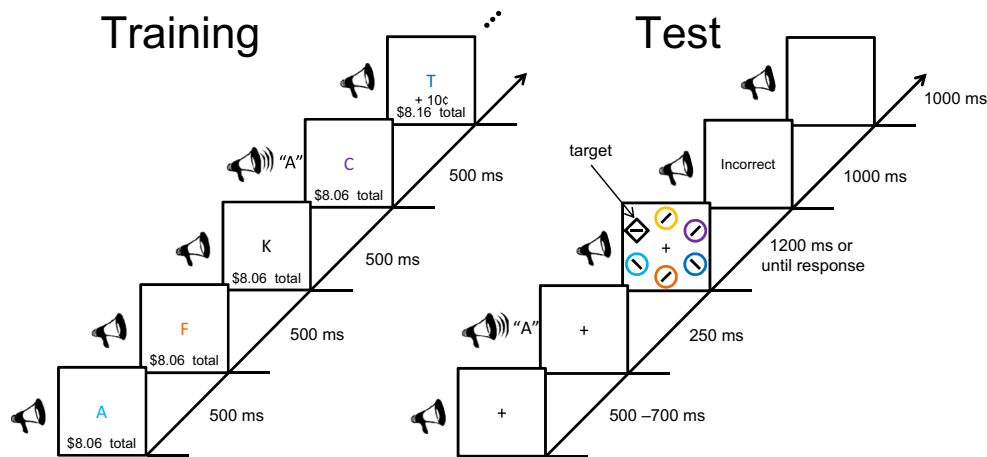
Fig. 1 Sequence and time course of trial events. **a** Training phase. Participants pressed the space bar any time an "A" or "Y" was presented auditorily while ignoring the visually presented letters. Detection of one target letter resulted in a comparatively high reward while detection of the other target letter resulted in a comparatively low reward. **b** Test phase. Participants searched for a shape singleton target (diamond among circles or circle among diamonds) and reported the orientation of the bar within the target as vertical or horizontal. Immediately prior to the onset of the search array, a sound from the training phase was played. The sound could be that of a former nontarget, the former low-reward target (low-value distractor), or the former high-reward target (high-value distractor)

played over the computer speakers. The spoken letter could be an "A," "Y," "X," or "H." The spoken letters were taken from the male voice stimulus set used by Shomstein and Yantis (2004) in an auditory attention task. The duration of each spoken letter was 240 ms, with 10 ms of silence added to the end of the sound file (for a total play time of 250 ms). A bank total reflecting earnings in the task was always visible 5.2° below the visually-presented letters, , centered along the x-axis in white 40-point font. Feedback concerning task performance appeared periodically between the visually presented letters and the bank total, also centered along the x-axis in white 40-point font.

Design For the visually presented letters, letter color and letter identity were randomly selected from the respective stimulus sets with the rule that no color or identity could repeat on consecutive letter displays. Each of the four spoken letters were presented equally often, the order of which was randomized within each block of trials. Therefore, the correspondence between the identity of the spoken letter and the concurrent visually presented letter was at chance (1/16). Each participant listened for the same two target letters: "A" and "Y." One target letter was assigned a high reward value while the other was assigned a low reward value that was awarded every time the target was correctly identified. Which target letter served as the high-reward target was counterbalanced across participants.

Procedure The training phase consisted of four blocks of trials, each of which comprised 610 letter-displays presented for a duration of 500 ms each. Each letter display contained a visually presented letter at the center of the screen. One hundred twenty of these letter displays also contained a spoken

letter, the duration of which lasted the first 250 ms of the letter display. Letter displays that included a spoken letter were randomly preceded by three, four, or five letter displays with no spoken letter. Once participants initiated the block by pressing the space bar, the entire letter-display set ran through to completion. After each block, participants were provided a brief rest period. There were no practice trials.

Participants were instructed to press the space bar as quickly as possible whenever they heard an "A" or "Y" spoken, which served as the target letters. They were informed that doing so would result in a monetary reward added to their total earnings. Participants were also instructed to withhold responding to the other two spoken letters and were informed that doing the opposite would result in a small amount of money being deducted from their total earnings. Participants were also informed that the money they earned from the first part of the experiment would serve as their compensation for completing the entire study. Participants earned 10¢ whenever they identified the high-reward target sound and 2¢ whenever they identified the low-reward target sound; they lost 5¢ for pressing the space bar in response to a nontarget sound. The instructions made no reference to the amounts of money that could be gained or lost, or any association between these amounts and the letter identities, which had to be learned from experience in the task. Following a depression of the space bar within 1,500 ms of the target sound presentation, participants received feedback indicating "+10¢" or "+ 2¢," while their bank total was simultaneously updated. If participants did not respond after 1,500 ms, the word "Miss!" was presented as feedback, and if participants responded within 1,500 ms of a nontarget sound, they received feedback indicating "- 5¢," while their bank total was simultaneously updated. Miss and loss feedback remained on the screen for one letter display,

and reward feedback remained on the screen for one letter display plus any letter displays occurring after a response was recorded that fell within the 1,500 ms response deadline. Note that reward was predicted only by the identity of the target sound and not by an associated motor response (which was the same for each target sound).

Test phase

Stimuli Each trial consisted of a fixation display, a search array, and (in the event that a correct response was not registered) a feedback display (see Fig. 1b). The fixation display contained a white fixation cross ($0.7° \times 0.7°$) presented in the center of the screen, and the search array consisted of the fixation cross surrounded by six colored shapes (each $2.4° \times 2.4°$) presented along an imaginary circle with a radius of $4.8°$. The six shapes comprising the search array consisted of either a diamond among circles or a circle among diamonds, and the target was defined as the unique shape. The color of each shape was drawn from the same set of six colors used during training (blue, cyan, pink, orange, yellow, white) without replacement on each trial. Different colors were included in the stimulus array to increase nontarget heterogeneity and thereby reduce the salience of the target (Duncan & Humphreys, 1989), maximizing the ability of the auditory distractors to effectively compete for attention. The feedback display, if presented, consisted of the words "Incorrect" or "Too Slow" in white 40-point font at the center of the display. All stimuli were presented on a black background.

Immediately preceding the search array, one of three sounds from the training phase was presented. The sounds used were the previous high-reward target (high-value distractor), the previous low-reward target (low-value distractor), and a previous nontarget (the sound "X").

Design One-third of the trials contained a high-value distractor, one-third a low-value distractor, and one-third a former nontarget distractor. For each of these trial types, the target was presented in each location equally often, the bar inside of which was equally often vertical and horizontal. Thus, in the test phase, the distractor sounds were completely irrelevant to the performance of the visual search task.

Procedure The experimenter started the test phase at the request of the participant, immediately following a self-paced break between the training and test phases. Participants were instructed to ignore the spoken letters and to focus on identifying the oriented bar within the unique shape. The test phase consisted of 144 trials, which were preceded by 20 practice trials that did not include the spoken letters.

Each trial began with a fixation display that was presented for a randomly varying interval of 500, 600, or 700 ms. Then, while the fixation display was still on screen, one of the three

sounds was presented over the speakers for the full 250 ms duration. Immediately after the sound was finished playing, the search array appeared and remained on screen until a response was made or 1,200 ms had elapsed, after which the trial timed out. Sounds were presented prior to the presentation of the stimulus array in order to ensure that the identity of the sound, which unfolds over time, was sufficiently processed before the target could be selected. The timing of the distractor sounds relative to the stimulus array was informed by the time required for an auditory target to elicit an attentional blink in the visual domain (Arnell & Jolicoeur, 1999).

Participants made a forced-choice target identification by pressing the "z" and the "m" keys for the vertically and horizontally orientated bars within the targets, respectively. Thus, the stimuli in the search array were linked to a different response rule than the auditory targets in the training phase, thereby minimizing contributions from response priming by the distractor sounds to RT. The search array was followed immediately by error feedback (the words "Incorrect" or "Too Slow") for 1,000 ms in the event that a correct response was not registered (this display was omitted following a correct response) and then by a blank 1,000 ms intertrial interval. No monetary rewards were given in the test phase. At the conclusion of the test phase, participants were paid the amount they had earned in the training phase.

Data analysis In the training phase, RT in target detection served as the primary measure of interest. Hit rate (percentage of targets detected) and false alarm rate (percentage of nontargets eliciting a response) were also computed. In the test phase, both responses that were incorrect and responses that exceeded the timeout limit were scored as errors; only correct responses were included in the analysis of RT. For both phases of the experiment, RTs more than 2.5 standard deviations above or below the mean of their respective condition for each participant were trimmed.

Results and discussion

Training phase Participants were significantly faster to detect a high-reward target (M = 560 ms) than a low-reward target (M = 591 ms), $t(25)$ = 4.67, $p < .001$, d = .92. This suggests that high-reward targets had greater attentional priority than low-reward targets. Because participants were not explicitly informed of the reward structure of the task, this attentional bias was the result of reward learning. Hit rate was very high and did not differ for high- and low-reward targets, $t(25)$ = 0.48, p = .645 (99.4 % and 99.3 %, respectively). False alarm rate was very low (0.8 %).

To examine the evolution of the observed attentional bias for the high-reward target over time, the RT difference between high- and low-reward targets was computed separately by block and compared across the four blocks of training. This

analyses revealed no main effect of block, $F(3, 75) = 1.29$, $p = .285$, ($M = 28, 37, 30$, and 41 ms across blocks). Although the attentional bias increased by 13 ms from the first to last block, this difference was not reliable, $t(25) = 1.61$, $p = .119$. An attentional bias for the high-reward target was evident as early as the first block, $t(25) = 4.04$, $p < .001$, $d = .79$, suggesting that the reward learning and consequent effects on attentional priority unfolded rapidly in this task.

Test phase A repeated-measures analysis of variance (ANOVA) on mean RT with distractor condition (former nontarget, low-value, high-value) as a factor revealed a main effect, $F(2, 50) = 5.59$, $p = .006$, $\eta_p^2 = .183$ (see Fig. 2). Participants were slower to report the target when a high-value distractor sound was emitted ($M = 700$ ms) compared to both a low-value distractor sound ($M = 684$ ms), $t(25) = 2.59$, $p = .016$, $d = .51$, and a former nontarget sound ($M = 681$ ms), $t(25) = 3.64$, $p = .001$, $d = .71$. The difference between the high- and low-value distractor conditions can only be explained in terms of relative value, as the actual sounds used were counterbalanced across these conditions. Accuracy did not differ by distractor condition, $F(2, 50) = 0.17$, $p = .841$. Accuracy was 84.0 %, 84.3 %, and 84.7 % across the former nontarget, low-value, and high-value distractor conditions, respectively.

Experiment 2

Experiment 1 provides evidence that sounds previously associated with high reward automatically capture attention in a manner that biases stimulus competition away from input arising from the visual system. Such value-dependent auditory distraction occurred when participants were actively searching for a visual target, consistent with cross-modal value-driven attentional capture. However, there are aspects of the experimental design that complicate interpretation of the observed distraction.

During training, the auditory targets were associated with a motor response (participants were free to choose which hand they responded to targets with). Responding to visual targets during the 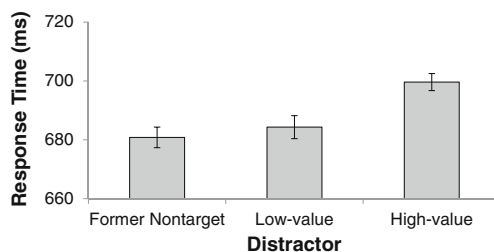test phase involved responding with each hand, depending on the orientation of the bar contained within the target, and thus required either the same or a different motor response than that previously associated with the high- and low-value distractor sounds. To the degree that the distractors elicited their previously associated motor response, the high-value distractor may have given rise to increased response conflict on some trials, which may have partly accounted for the observed slowing of RT in that condition.

Also, during the test phase of Experiment 1, the distractors were presented in advance of the stimulus array. This was done to ensure that the distractors were sufficiently processed to the degree that they would be able to effectively compete for attention with a visual target. However, although irrelevant to the visual search task itself, all distractors predicted the exact timing of the visual search array and participants may have therefore voluntarily attended to the distractors (in spite of the instruction to try to ignore them) in order to better prepare for the upcoming visual task. To the degree that this occurred, the cost in performance associated with the high-value distractor may be explained by delayed disengagement following voluntary attentional orienting rather than the initial capture of attention.

Finally, irrelevant visual letters were presented during the training phase in order to pair the selection of targets with the need to ignore visual input. Although maximizing the magnitude of possible attention effects, this creates ambiguity in the nature of the value-driven attentional bias. It is unclear the degree to which the value-driven bias observed in Experiment 1 reflects the selection of the auditory distractors vs the suppression of concurrent visual input.

To address these issues, Experiment 2 involved a conceptually similar approach with a simpler design. No visual letters were presented during training, and participants responded with different hands during the training phase (left hand) and test phase (right hand). The distractor sounds were played simultaneously with the onset of the stimulus array, making them no more predictive of the stimulus array than the onset of the visual stimuli itself. Finally, the different colors were removed from the stimulus array at test (all stimuli were white) in order to enhance the visual salience of the target (Duncan & Humphreys, 1989) and thereby examine the robustness of value-driven auditory attentional capture.

Method

Participants Thirty-two new participants were recruited from the Johns Hopkins University community. All reported normal or corrected-to-normal visual acuity and normal color vision. Data for three participants were replaced: one for poor task performance (accuracy < 60 %) and two using the same outlier criterion as Experiment 1.

Fig. 2 Mean response time by distractor condition in the test phase of Experiment 1. Error bars reflect the within-subjects *SEM*

Apparatus, stimuli, and procedure The apparatus, stimuli, and procedure were identical to those of Experiment 1 with the following exceptions. All visual letters were eliminated from the training phase, leaving only the reward information on the computer screen. Participants reported auditory targets using the "z" key with their left hand during the training phase. During the test phase, they responded using the "n" and "m" keys with the index and middle finger of their right hand for targets containing a vertical and horizontal bar, respectively. All of the shapes were white during the test phase, and the distractor sounds began playing simultaneous with the onset of the stimulus array on each trial.

Results

Training phase Participants were again significantly faster to detect a high-reward target ($M = 589$ ms) than a low-reward target ($M = 619$ ms), $t(31) = 4.95$, $p < .001$, $d = .87$, suggesting that high-reward targets had greater attentional priority. Hit rate was very high and did not differ for high- and low-reward targets, $t(31) = 1.44$, $p = .161$ (99.6 % and 99.4 %, respectively). False alarm rate was very low (0.7 %).

As in Experiment 1, I examined the evolution of the observed attentional bias for the high-reward target across blocks. There was no main effect of block, $F(3, 93) = 2.12$, $p = .103$, ($M = 18$, 30, 34, and 39 ms across blocks). Comparing the attentional bias between the first and last block showed a reliable difference in which this bias became larger over the course of training, $t(31) = 2.25$, $p = .032$, $d = .40$. An attentional bias for the high-reward target was again evident as early as the first block of training, $t(31) = 2.22$, $p = .034$, $d = .39$.

Test phase Planned comparisons, based on the results of Experiment 1, replicated the finding that participants were significantly slower to report the target when a high-value distractor sound was emitted ($M = 683$ ms) compared to a low-value distractor sound ($M = 673$ ms), $t(31) = 2.31$, $p = .028$, $d = .41$ (see Fig. 3). The difference in RT between the high-value and former nontarget ($M = 681$ ms) distractor conditions, however, was not reliable, $t(31) = 0.39$, $p = .701$, nor was there a reliable difference between the low-value and former nontarget distractor conditions, $t(31) = -1.70$, $p = .099$.
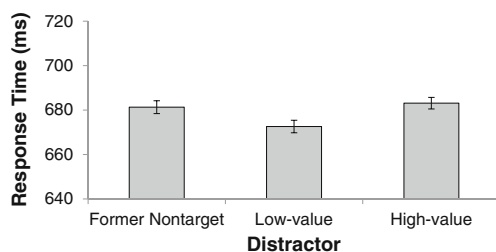

Fig. 3 Mean response time by distractor condition in the test phase of Experiment 2. Error bars reflect the within-subjects *SEM*

A similar pattern was observed in accuracy, which was lower for high-value distractor trials compared to low-value distractor trials, $t(31) = 2.74$, $p = .010$, $d = .48$, but did not differ between former nontarget distractor trials and either high-value or low-value distractor trials, $t(31) = 0.75$, $p = .458$, and, $t(31) = -1.53$, $p = .137$, respectively. Accuracy was 89.4 %, 90.8 %, and 88.6 % across the former nontarget, low-value, and high-value distractor conditions, respectively.

Between-experiments comparison The slowing of RT associated with the high-value distractor when compared to the low-value distractor condition did not significantly differ between experiments (M difference = 5 ms), $t(56) = 0.51$, $p = .512$. However, the slowing of RT associated with the high-value distractor when compared to the former nontarget distractor condition was significantly greater in Experiment 1 (M difference = 17 ms), $t(56) = 2.41$, $p = .019$, $d = .64$.

Discussion

Experiment 2 replicates the critical finding from Experiment 1 in that the same auditory stimulus is more distracting when it was previously associated with high versus low reward. This difference in performance was observed under conditions in which the potential for distractor-evoked response competition and the predictiveness of the distractor were minimized and where the training task emphasized only the selection of auditory information. The magnitude of this value-driven impairment in performance was similar to that observed in Experiment 1.

In contrast, the performance impairment relative to nontarget distractor trials was not replicated in Experiment 2. Although unexpected, the pattern of performance observed in Experiment 2 is not without precedent. Anderson et al. (2012) examined whether learned attentional priorities for a reward-predictive color could transfer to a new experimental task (visual search to a flankers task), and observed greater flanker compatibility effects for flankers rendered in the high-value compared to the low-value color, with the compatibility effects for former nontarget colored flankers falling nonsignificantly between these two conditions. A follow-up experiment that involved otherwise comparable but unrewarded training demonstrated that former nontarget colors received greater attentional priority than former target colors, which was interpreted as a bias to prioritize less familiar stimuli (see Johnston, Hawley, Plewe, Elliott, & DeWitt, 1990) in a new situation. Because the training and test phases differed to a greater extent in Experiment 2 (both phases involved the presentation of color stimuli in Experiment 1 whereas visual stimuli were presented only in the test phase in Experiment 2, the hand used to input responses differed between phases in Experiment 2 but not Experiment 1), this might explain the reduced attentional priority of the previously reward-

associated auditory distractors relative to the former nontarget distractor condition. An alternative interpretation of the lack of difference in performance between the high-value and former nontarget distractor conditions in Experiment 2 is that selection of the low-value distractor was inhibited. Most critically, however, both experiments demonstrate a modulation of attentional processing attributable only to the associated value of a former target, which argues in favor of value-dependence in cross-modal attentional capture by auditory stimuli. The ambiguity in the data caused by the former nontarget distractor condition is addressed in Experiment 3.

Experiment 3

Experiment 3 tests the hypothesis that the elevated RTs to nontarget distractors in Experiment 2 was influenced by an attentional bias for less familiar stimuli, similar to that observed in Anderson et al. (2012). To that end, Experiment 3 repeated the procedures of Experiment 2 but without reward feedback during training. If attention is indeed biased towards less familiar stimuli, RT to report the shape target should be slowed by former nontarget distractors compared to former target distractors that were never associated with reward during training.

Method

Participants Sixteen new participants were recruited from the Johns Hopkins University community. All reported normal or corrected-to-normal visual acuity and normal color vision. Data for one participant was replaced using the same outlier criterion as Experiment 1.

Apparatus, stimuli, and procedure The apparatus, stimuli, and procedure were identical to those of Experiment 2 with the exception that the reward feedback was removed. No bank total was presented on the screen during training and the word "Correct!" replaced the monetary increment that appeared after each correct response.

Data analysis Because there was no value difference between the two target sounds during training, data from distractor trials containing these stimuli were collapsed into a former target distractor condition as in Anderson et al. (2012). Otherwise, the data analysis procedures were the same as in the prior experiments.

Results

Training phase Mean RT during the training phase was 623 ms, hit rate was 98.7 %, and false alarm rate was 2.0 %.

Test phase Mean RT was significantly slower on former non-target distractor trials (678 ms) compared to former target distractor trials (667 ms), $t(15) = 2.37$, $p = .032$, $d = 0.59$. Accuracy did not differ between the former nontarget and former target distractor conditions (88.4 % and 87.4 %, respectively), $t(15) = 0.92$, $p = .372$.

Discussion

The results of Experiment 3 are consistent with an attentional bias for less familiar stimuli (Johnston et al., 1990) in a new task, replicating the pattern observed in Anderson et al. (2012) using auditory distractors. Because the assignment of specific sounds to conditions was only counterbalanced between the high- and low-value conditions and not between these and the nontarget condition, an alternative possibility is that the specific nontarget sound chosen had a higher intrinsic salience than the other sounds. In either case, however, the results of Experiment 3 suggest that the elevated RTs to the former nontarget distractors in Experiment 2 reflect elevated priority for these stimuli rather than suppression of the low-value distractor.

General discussion

Attentional selection is influenced by bottom-up physical salience (e.g., Theeuwes, 1992), top-down goals (e.g., Folk et al., 1992), and learned value (e.g., Anderson et al., 2011b). Although there is now ample evidence that principles of salience-driven and goal-directed attention broadly apply to different sensory modalities and bias cross-modal stimulus competition (e.g., Arnell & Jolicoeur, 1999; Dalton & Lavie, 2004; Dalton & Spence, 2007; Spence, 2010; Spence & Driver, 1994, 1997), evidence for value-driven attention has been restricted entirely to the domain of vision.

It cannot be straightforwardly assumed that the principle of value-driven attention can be similarly extended to sensory modalities beyond vision. For example, the neural correlates of value-driven attention to visual stimuli are confined largely to the visual system, encompassing extrastriate cortex, intraparietal sulcus/lateral intraparietal area, and the visual cortico-striatal loop (e.g., Anderson, Laurent, & Yantis, 2014; Hickey & Peelen, 2015; Krebs, Boehler, Egner, & Woldorff, 2011; Peck et al., 2009; Qi et al., 2013). The present study provides direct evidence that reward learning has a broad impact on attention that extends beyond the visual system. Value-driven attentional capture by an auditory stimulus can even compete with the representation of a stimulus in vision, reflecting a biasing signal that can divert attention from one sensory modality to another.

The present study connects the principle of value-driven attention to the extensive literature on auditory attention and cross-modal attention (e.g., Rhodes, 1987; Spence, 2010; Wu,

Weissman, Roberts, & Woldorff, 2007). My findings suggest that value-driven attention reflects, at least in part, the allocation of a domain-general processing resource capable of biasing information processing in favor of input originating from different sensory modalities. Value-driven auditory attention provides a cognitive mechanism that could be useful for explaining a variety of biases and asymmetries in auditory processing, such as the attention-capturing quality of hearing one's own name spoken (e.g., Moray, 1959). The present study also sheds light on the kinds of sensory experiences that can interfere with goal-directed processing, with implications for our understanding of vulnerability to addiction relapses (see Anderson et al., 2013).

References

Anderson, B.A. (2013). A value-driven mechanism of attentional selection. *Journal of Vision, 13*(3),7, 1–16.

Anderson, B. A. (2015a). Value-driven attentional capture is modulated by spatial context. *Visual Cognition, 23,* 67–81.

Anderson, B. A. (2015b). Value-driven attentional priority is context specific. *Psychonomic Bulletin and Review, 22,* 750–756.

Anderson, B. A., Faulkner, M. L., Rilee, J. J., Yantis, S., & Marvel, C. L. (2013). Attentional bias for non-drug reward is magnified in addiction. *Experimental and Clinical Psychopharmacology, 21,* 499–506.

Anderson, B. A., Laurent, P. A., & Yantis, S. (2011a). Learned value magnifies salience-based attentional capture. *PLoS ONE, 6*(11), e27926.

Anderson, B. A., Laurent, P. A., & Yantis, S. (2011b). Value-driven attentional capture. *Proceedings of the National Academy of Sciences of the United States of America, 108,* 10367–10371.

Anderson, B. A., Laurent, P. A., & Yantis, S. (2012). Generalization of value-based attentional priority. *Visual Cognition, 20,* 647–658.

Anderson, B. A., Laurent, P. A., & Yantis, S. (2014a). Value-driven attentional priority signals in human basal ganglia and visual cortex. *Brain Research, 1587,* 88–96.

Anderson, B. A., Leal, S. L., Hall, M. G., Yassa, M. A., & Yantis, S. (2014b). The attribution of value-based attentional priority in individuals with depressive symptoms. *Cognitive, Affective, & Behavioral Neuroscience, 14,* 1221–1227.

Anderson, B. A., & Yantis, S. (2012). Value-driven attentional and oculomotor capture during goal-directed, unconstrained viewing. *Attention, Perception, & Psychophysics, 74,* 1644–1653.

Anderson, B. A., & Yantis, S. (2013). Persistence of value-driven attentional capture. *Journal of Experimental Psychology: Human Perception and Performance, 39,* 6–9.

Arnell, K. M., & Jolicoeur, P. (1999). The attentional blink across stimulus modalities: Evidence for central processing limitations. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 630–648.

Bacon, W. F., & Egeth, H. E. (1994). Overriding stimulus-driven attentional capture. *Perception & Psychophysics, 55,* 485–496.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10,* 433–436.

Chelazzi, L., Estocinova, J., Calletti, R., Lo Gerfo, E., Sani, I., Della Libera, C., & Santandrea, E. (2014). Altering spatial priority maps via reward-based learning. *Journal of Neuroscience, 34,* 8594–8604.

Dalton, P., & Lavie, N. (2004). Auditory attentional capture: Effects of singleton distractor sounds. *Journal of Experimental Psychology: Human Perception and Performance, 30,* 180–193.

Dalton, P., & Lavie, N. (2007). Overriding auditory attentional capture. *Perception & Psychophysics, 69,* 162–171.

Dalton, P., & Spence, C. (2007). Attentional capture in serial audiovisual search tasks. *Perception & Psychophysics, 69,* 422–438.

Della Libera, C., & Chelazzi, L. (2006). Visual selective attention and the effects of monetary reward. *Psychological Science, 17,* 222–227.

Della Libera, C., & Chelazzi, L. (2009). Learning to attend and to ignore is a matter of gains and losses. *Psychological Science, 20,* 778–784.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18,* 193–222.

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review, 96,* 433–458.

Duncan, J., Martens, S., & Ward, R. (1997). Restricted attentional capacity within but not between sensory modalities. *Nature, 387,* 808–810.

Failing, M.F., & Theeuwes, J. (2014). Exogenous visual orienting by reward. *Journal of Vision, 14*(5), 6, 1–9.

Folk, C. L., & Anderson, B. A. (2010). Target-uncertainty effects in attentional capture: Color-singleton set or multiple attentional control settings? *Psychonomic Bulletin and Review, 17,* 421–426.

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 1030–1044.

Hickey, C., Chelazzi, L., & Theeuwes, J. (2010). Reward changes salience in human vision via the anterior cingulate. *Journal of Neuroscience, 30,* 11096–11103.

Hickey, C., & Peelen, M. V. (2015). Neural mechanisms of incentive salience in naturalistic human vision. *Neuron, 85,* 512–518.

Johnston, W. A., Hawley, K. J., Plewe, S. H., Elliott, J. M. G., & DeWitt, M. J. (1990). Attention capture by novel stimuli. *Journal of Experimental Psychology: General, 119,* 397–411.

Krebs, R. M., Boehler, C. N., Egner, T., & Woldorff, M. G. (2011). The neural underpinnings of how reward associations can both guide and misguide attention. *Journal of Neuroscience, 31,* 9752–9759.

Laurent, P. A., Hall, M. G., Anderson, B. A., & Yantis, S. (2015). Valuable orientations capture attention. *Visual Cognition, 23,* 133–146.

Le Pelley, M. E., Pearson, D., Griffiths, O., & Beesley, T. (2015). When goals conflict with values: Counterproductive attentional and oculomotor capture by reward-related stimuli. *Journal of Experimental Psychology: General, 144,* 158–171.

Lee, J., & Shomstein, S. (2014). Reward-based transfer from bottom-up to top-down search tasks. *Psychological Science, 25,* 466–475.

Miranda, A. T., & Palmer, E. M. (2014). Intrinsic motivation and attentional capture from gamelike features in a visual search task. *Behavior Research Methods, 46,* 159–172.

Moray, N. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology, 11,* 56–60.

Pantoja, J., Ribeiro, S., Wiest, M., Soares, E., Gervasoni, D., Lemos, N. A., & Nicolelis, M. A. (2007). Neuronal activity in the primary somatosensory thalamocortical loop is modulated by reward contingency during tactile discrimination. *Journal of Neuroscience, 27,* 10608–1062010.

Peck, C. J., Jangraw, D. C., Suzuki, M., Efem, R., & Gottlieb, J. (2009). Reward modulates attention independently of action value in posterior parietal cortex. *Journal of Neuroscience, 29,* 11182–11191.

Pooresmaeili, A., FitzGerald, T. H. B., Bach, D. R., Toelch, U., Ostendorf, F., & Dolan, R. J. (2014). Cross-modal effects of value on perceptual acuity and stimulus encoding. *Proceedings of the National Academy of Sciences of the United States of America, 111,* 15244–15249.

Qi, S., Zeng, Q., Ding, C., & Li, H. (2013). Neural correlates of reward-driven attentional capture in visual search. *Brain Research, 1532,* 32–43.

Raymond, J. E., & O'Brien, J. L. (2009). Selective visual attention and motivation: The consequences of value learning in an attentional blink task. *Psychological Science, 20,* 981–988.

Rhodes, G. (1987). Auditory attention and the representation of spatial information. *Perception & Psychophysics, 42,* 1–14.

Sali, A. W., Anderson, B. A., & Yantis, S. (2014). The role of reward prediction in the control of attention. *Journal of Experimental Psychology: Human Perception and Performance, 40,* 1654–1664.

Shomstein, S., & Yantis, S. (2004). Control of attention shifts between vision and audition in human cortex. *Journal of Neuroscience, 24,* 10702–10706.

Spence, C. (2010). Crossmodal spatial attention. *Annals of the New York Academy of Sciences, 1191,* 182–200.

Spence, C., & Driver, J. (1994). Covert spatial orienting in audition: Exogenous and endogenous mechanisms. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 555–574.

Spence, C., & Driver, J. (1997). Audiovisual links in exogenous covert spatial orienting. *Perception & Psychophysics, 59,* 1–22.

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics, 51,* 599–606.

Theeuwes, J., & Belopolsky, A. V. (2012). Reward grabs the eye: Oculomotor capture by rewarding stimuli. *Vision Research, 74,* 80–85.

Wu, C.-T., Weissman, D. H., Roberts, K. C., & Woldorff, M. G. (2007). The neural circuitry underlying the executive control of auditory spatial attention. *Brain Research, 1134,* 187–198.

Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 350–374.

Yantis, S., Schwarzbach, J., Serences, J. T., Carlson, R. L., Steinmetz, M. A., Pekar, J. J., & Courtney, S. M. (2002). Transient neural activity in human parietal cortex during spatial attention shifts. *Nature Neuroscience, 5,* 995–1002.