

RUNNING HEAD: Statistical Learning of Attentional Set

In press at Journal of Experimental Psychology: Human Perception Performance

Establishment of an attentional set via statistical learning

Joshua D. Cosman¹
and
Shaun P. Vecera²

¹Department of Psychology, Vanderbilt Vision Research Center, Vanderbilt University

²Departments of Psychology and Neuroscience, University of Iowa

Address correspondence to Joshua Cosman at:

Vanderbilt University
Department of Psychology
Wilson Hall
PMB 407817
2301 Vanderbilt Place
Nashville, TN 37240-7817
joshua.d.cosman@vanderbilt.edu

WORD COUNT: 2786

Keywords: Visual Attention; Attentional Capture; Statistical Learning

Abstract

The ability overcome attentional capture and attend goal-relevant information is typically viewed as a volitional, effortful process that relies on the maintenance of current task priorities or “attentional sets” in working memory. However, the visual system possesses statistical learning mechanisms that can incidentally encode probabilistic associations between goal-relevant objects and the attributes likely to define them. Thus, it is possible that statistical learning may contribute to the establishment of a given attentional set and modulate the effects of attentional capture. Here we provide evidence for such a mechanism, showing that implicitly learned associations between a search target and its likely color directly influence the ability of a salient color precue to capture attention in a classic attentional capture task. This indicates a novel role for statistical learning in the modulation of attentional capture, and emphasizes the role that this learning may play in goal-directed attentional control more generally.

Establishment of an attentional set via statistical learning

Salient information that is unrelated to our goals sometimes captures our attention, distracting us from the task at hand. Our ability to exercise goal-directed control over attentional capture depends critically on the “attentional set” adopted by an individual, which consists of information regarding our immediate search goals (Folk et al., 1992; Bacon & Egeth, 1994). Although the precise representational basis of an attentional set is unknown, theories of attention propose a central role for working memory processes in the on-line representation of task goals, with the active maintenance of information regarding ongoing task priorities being used to flexibly guide attention on a moment to moment basis (Desimone & Duncan, 1995; Bundesen, 1990; Bundesen, Habekost, & Kyllingsbaek, 2005). Such a view has been influential in describing goal-related influences on attentional control, gaining support from a number of studies demonstrating that discrete information (e.g., colors, shapes, spatial locations) explicitly rehearsed in working memory can bias the allocation of attention (Downing, 2000; Awh, Jonides, & Reuter-Lorenz, 1998; Soto et al., 2005; Olivers et al., 2006; Woodman & Luck, 2007; Cosman & Vecera, 2011).

Despite this emphasis on explicit working memory processes in the operation of goal-directed attention, information that is not explicitly represented but nevertheless bears on effective execution of task goals may be used to increase the efficiency of goal-directed attentional control. Specifically, the visual system possesses statistical learning mechanisms that allow the acquisition of knowledge regarding the featural, temporal, or spatial characteristics likely to define objects (i.e., ‘visual statistical learning’ Fiser & Aslin, 2001, 2002; Turk-Browne et al., 2008). This learning can influence the

deployment of attention even when probability information is acquired incidentally and is not explicitly represented (e.g., Chun & Jiang, 1998, 1999; Ryan, Altoff, Whitlow, & Cohen, 2000; Geng & Behrmann, 2005), indicating that memory systems responsible for visual statistical learning can play a direct role in attentional control.

However, the extent to which incidental visual statistics contribute to the formation of an attentional set and influence capture is unknown, primarily because previous studies have confounded explicit task goals (e.g., “search for the red target”) with the features that are likely to define goal-relevant information (e.g., the color red), obscuring possible influences of this learning on task performance. The primary goal of the current work was to dissociate explicit task goals from the features likely to define goal-relevant information, in order to determine whether statistical learning of target-defining features is itself sufficient to drive the emergence of a given attentional set.

We had observers complete a training task where they performed a spatial cueing task similar to that used to provide evidence that explicit attentional sets can influence attentional capture (Folk et al., 1992), with a handful of manipulations that allowed us to examine the contribution of statistical learning to the establishment of a given attentional set (Figure 1). In our task, observers were asked to search for and report the identity of a target letter (B vs. H) that could appear in one of two colors (red or green) on a given trial. Critically, we introduced an asymmetry in the probability that the target would appear in one color or the other, with the target drawn in one color on 80% of trials and in the other color on 20% of trials. Observers were then encouraged to take a break, the experimenter set up the testing session, and approximately five to ten minutes later each observer completed a testing task that was identical to that used during training, except

that the target-color asymmetry was removed. This allowed a pure assessment of statistical learning without influence from possible color/probability confounds present during the training session, which may reflect a mixed influence of statistical learning and intertrial priming (see, e.g, Folk & Remington, 2008; Belopolsky, Schreij, & Theeuwes, 2010).

Of primary interest was whether, during the testing session, cues matching the color that had been more likely to define the target during the training session would lead to larger capture effects than those matching the less likely target color, producing what would be considered a “classic” contingent capture effect (Folk et al., 1992),¹ and indicating that observers had learned the contingency present in the training task. In addition, at the end of the experiment observers completed a questionnaire probing their awareness of the probability manipulation present in the training session, allowing us to examine the influence of explicit knowledge of target-color relationships on control over attentional capture.

Method

Observers. Observers were 17 University of Iowa undergraduates who participated for course credit. All had normal or corrected to normal vision and were not color blind.

¹ Because we made a handful of minor alterations to the traditional contingent capture task (Folk et al., 1992), we performed an experiment to verify that our stimuli could produce expected contingent capture effects given an explicit set for target color. Ten observers were instructed to explicitly search for a target of a particular color (red or green, counterbalanced across observers) and report its identity. Results indicated a typical contingent capture effect, with cues drawn in a color matching the observers’ explicit set producing a significant cueing effect (20ms; $t(9) = 2.65$, $p = .02$), whereas non-matching cues did not (1ms; $t < 1$, $p = \text{n.s.}$).

Apparatus. Stimuli were presented on a 15” CRT monitor powered by a Macintosh Mini computer, using MATLAB and the Psychophysics Toolbox (Brainard, 1997).

Stimuli & Design. Observers sat approximately 65 cm from the screen and viewed displays resembling those in Figure 1. The fixation display consisted of 2 placeholder boxes measuring $1.4^\circ \times 1.4^\circ$, positioned to the left and right of fixation. The distance from fixation to the center of each placeholder box was 5.2° . The placeholder boxes were light gray on a black background. Cues consisted of a single set of 4 dots (radius $.21^\circ$) centered on the edges of a placeholder box, with each dot positioned $.46^\circ$ peripheral to the side of the placeholder. The cues were spatially non-predictive and were equally likely to be presented in either red or green, selected pseudorandomly on each trial. A single target symbol was presented on each trial, and was either a ‘B’ or an ‘H’ drawn in 56 pt Helvetica bold font, with identity chosen pseudorandomly on each trial. The color of the target could be either red or green.

During the training session, for half of the observers the target appeared in red on 80% of trials and in green on the remaining 20% of trials, with this asymmetry being reversed for the other half of the observers. Concurrent with the target, a non-target symbol was presented in the remaining placeholder box, with non-target identities being drawn randomly from a set of letters consisting of K,L,T,V, each of which were also drawn in a 56 pt Helvetica bold font, and which were equally likely to appear in either red or green on a given trial. Following the training session observers completed a testing session in which the task was identical, but the color asymmetry was removed.

Procedure. Observers completed the experiment in a single session lasting approximately 40 minutes, and were instructed to search for a target letter while trying to ignore the uninformative precue. On each trial, a fixation display was presented for 1000ms, followed by a single spatially non-predictive red or green precue for 50ms, and then by a 100ms ISI. Directly following this, the search display was presented for 50ms (producing a cue-target SOA of 150ms). The duration from the time of cue onset to the time of target offset was 200ms, a duration short enough to preclude eye movements to the cue or target locations. The fixation display remained until observers made a response using their index finger of each hand to press either the 'Z' or 'M' keys, with target-response mappings counterbalanced across observers.

The training session consisted of 12 blocks of 40 trials (480 total trials) and the testing session consisted of four blocks of 40 trials (160 total trials). In order to assess awareness of the target-color asymmetry present during training, directly following completion of the testing task observers were given a short questionnaire to probe their awareness of the color manipulations present during the training session. They were asked 1) if they performed the task using any specific strategies and 2) whether they noticed any regularities in the colors used in the training task, allowing us to probe the whether observers had employed an explicit set for color the likely target color. Following these open-ended questions, observers were informed of the color manipulation and given a forced choice task and asked to select which color, red or green, was more likely to define the target during training. By comparing the magnitude of capture effects across individuals who answered this question correctly/incorrectly, we

were able to quantitatively assess possible influences of latent explicit knowledge of the color asymmetry on task performance.

Results

Incorrect trials and trials with reaction times (RTs) greater than 1000ms were excluded from further analysis, with outlier trimming resulting in a removal of approximately 4% of the RT data from the training session and 2% of the RT data from the testing session. Observers' overall mean correct RT and error rate data for the training and testing sessions appear in Figures 2 and 3, respectively. For each session, reaction time (RT) and error rate data were entered into a two-factor ANOVA with cue color (match more likely (80%) target color vs. match less likely (20%) target color) and cue validity (valid vs. invalid) as factors.

Training Data. For RTs (Figure 2), we observed a significant main effect of cue validity, $F(1,16) = 19.3$, $p < .001$, $\eta^2 = .55$, indicating faster responses to validly cued than invalidly cued targets. The main effect of cue color was not significant, $F(1,16) = 1.7$, $p = .21$. However, we did find a significant interaction between cue color and cue validity, $F(1,16) = 10.5$, $p < .01$, $\eta^2 = .40$, with larger cueing effects for cues that matched the more likely target color (20ms) than those matching the less likely target color (8ms). Analysis of error rates showed no significant main effects or interactions.

Testing Data. For RTs (Figure 3), we observed a significant main effect of cue validity, $F(1,16) = 28.9$, $p < .001$, $\eta^2 = .64$, indicating faster responses to validly cued than invalidly cued targets. The main effect of cue color was not significant, $F < 1$, $p = n.s.$ Importantly, we again observed a significant interaction between cue color and cue validity, $F(1,16) = 17.1$, $p = .001$, $\eta^2 = .52$, with the magnitude of the cueing effect

remaining larger for the cue that matched the more likely target color (27ms) than the cue that matched the less likely target color (11ms) even in the absence of the target-color asymmetry.

In order to examine whether this effect was sensitive to the change in environmental statistics introduced during the testing session, we examined the magnitude of cueing effects for cues matching both the more and less likely target color across testing blocks using Bonferroni corrected t-tests. Cueing effects were significantly larger for cues matching the more likely target color in the first block of the testing session, $t(16)=2.5$, $p=.02$, and marginally so in the second block, $t(16)=1.4$, $p=.17$, but not subsequent blocks, all $t_s < 1$, *n.s.*, indicating gradual re-adaptation to the change in statistics. As in the training session, analysis of error rates showed no significant main effects or interactions.

Awareness analysis. Only two observers accurately described the color contingency on the open-ended questions, and none reported using an explicit set for color, suggesting a general lack of awareness of the color asymmetry. Ten of the seventeen observers answered the forced choice question correctly, and we quantitatively assessed possible latent knowledge of the color asymmetry on statistical learning by entering cueing effects (invalid minus valid RTs) into a mixed-model ANOVA with forced choice response accuracy (correct response vs. incorrect response) and cue color (matched more likely target color vs. matched less likely target color) as factors. Critically, there was no interaction between response accuracy and cue color, $F(1, 16) < 1$, *n.s.*, with significantly larger cueing effects for cues matching the more likely target color in both groups (correct response group, 29 ms vs. 11 ms, $t(9) = 2.3$, $p = .04$;

incorrect response group, 24 ms vs. 1 ms, $t(6) = 2.4, p = .03$). Taken together, this suggests that the contingent capture effect observed here does not depend on an explicit representation of features related to the target of search.

Discussion

We have shown that feature-based attentional sets can arise more or less automatically, with incidental exposure to visual statistics being sufficient to drive the emergence of feature-based attentional sets. That these effects were observed even though the colors used in the task did not uniquely specify the task-relevant target (i.e., salient precues and non-targets were drawn in the same colors as the target item) seems to suggest that the influence of incidental learning on attention is confined to task relevant stimuli, a notion consistent with previous work (Chun & Jiang, 2001; Turk-Browne et al., 2005).

Our results complement recent studies demonstrating that intertrial priming mechanisms can exert a similar incidental influence on the establishment of an attentional set (Folk & Remington, 2008; Belopolsky et al., 2010), suggesting that although explicit rehearsal of discrete information in working memory can guide visual attention and modulate capture, such rehearsal is not necessary for highly specific and effective attentional sets to arise. Although our results are similar to those seen in the above studies our results differ in time course from traditional intertrial feature-priming effects; typical intertrial effects last 5-8 trials (Maljkovic & Nakayama, 1994) whereas our effects appear to last 50 or more trials following the removal of the target-color asymmetry, gradually returning to baseline in the absence of the predictive color-target relationship. However, both lines of work argue that the implementation of an attentional set may reflect the

attention system's ability to adapt to regularities in the environment, optimizing task performance regardless of (or possibly in spite of) an individual's explicit goals and it is possible that statistical learning and feature priming share a mechanistic basis (Mozer, Shettel, & Vecera, 2006).

This interpretation is in line with recent suggestions that goal-directed attentional control relies heavily on past experience, both in the short-term (e.g., via intertrial priming effects) and longer-term (e.g., via semantic or episodic memory; see Awh, Belopolsky, & Theeuwes, 2012; Hutchinson & Turk-Browne, 2012 for reviews). Most relevant to the current work, a number of studies have shown that task-specific learning can lead observers to implicitly develop attentional sets that do not precisely match those that they report to use explicitly (see, e.g., Leber & Egeth, 2006; Leber, Kawahara & Gabari, 2008; Kawahara, 2010). As a result, it has been suggested that experience with specific attributes of a given task may be a critical factor determining the emergence and effectiveness of a given attentional set, irrespective of whether individuals are aware that this learning has influenced attentional control (Vaterott & Vecera, 2012; Cosman & Vecera, 2012, in press). In addition, our results are consistent with recent demonstrations that learned associations between a given stimulus feature and its reward value can modify attentional capture in an automatic manner (Anderson, Laurent, & Yantis, 2011). However, the current results suggest that explicit reward is not always required for driving feature-based learning effects on attentional capture, with these effects arising due to learned associations between a goal-relevant stimulus and its defining features in some situations.

Finally, although the results of the current experiment bear a close resemblance to those observed in traditional statistical learning tasks (e.g., they both arise incidentally on the basis of environmental statistics), the task used here differs in an important way from tasks typically used to study visual statistical learning. Specifically, whereas in our task observers learned the relationship between objects and their associated features (an intra-object form of statistical learning), traditional visual statistical learning tasks focus primarily on the learning of relationships between different objects (inter-object statistical learning; Fiser & Aslin, 2001, 2002; but see Turk-Browne et al, 2008). However, it has been argued that the multitude forms of incidental learning likely share a common mechanistic basis (Perruchet & Pacton, 2006), possibly relying on general purpose relational memory systems responsible for both intra- and inter-object binding (Diana, Yonelinas, & Ranganath, 2007; Konkel & Cohen, 2009).

Regardless of the precise mechanisms, the current work demonstrates a critical role for statistical learning in determining attentional capture, and a better understanding of how statistical learning systems interact with on-line control processes is critical to understanding the operation of attentional control more generally. By including information regarding regularities in the environment, attentional sets may be more 'holistic' than previously considered, representing the confluence of explicit goals and incidentally acquired representations of the larger context in which a task is performed. Such a mechanism would allow the visual system to ease the burden on discrete, capacity-limited working memory representations by shifting control to longer-term distributed representations that include information regarding relationships between objects and their likely features or spatial locations over time.

References

- Awh, E., Belopolsky, A.V., Theeuwes, J. (2012). Top-down versus bottom-up attentional control: a failed theoretical dichotomy. *Trends in Cognitive Science*, 16, 437-43.
- Awh, E., Jonides, J., & Reuter-Lorenz, P. A. (1998). Rehearsal in spatial working memory. *Journal of Experimental Psychology. Human Perception and Performance*, 24, 780-90.
- Bacon, W. F., & Egeth, H. E. (1994). Overriding stimulus-driven attentional capture. *Perception & Psychophysics*, 55, 485-96.
- Belopolsky, A. V., Schreij, D., & Theeuwes, J. (2010). What is top-down about contingent capture? *Attention, Perception, & Psychophysics*, 72, 326.
- Bundesen, C. (1990). A theory of visual attention. *Psychological review*, 97, 523-47.
- Bundesen, C., Habekost, T., & Kyllingsbaek, S. (2005). A neural theory of visual attention: bridging cognition and neurophysiology. *Psychological Review*, 112, 291-328.
- Chun, M M, & Jiang, Y. (1998). Contextual cueing: implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology*, 36, 28-71.
- Chun, M.M., & Jiang, Y. (1999). Top-down Attentional Guidance Based on Implicit Learning of Visual Covariation. *Psychological Science*, 10, 360-65.
- Cosman, J. D., & Vecera, S. P. (2011). The contents of visual working memory reduce uncertainty during visual search. *Attention, Perception & Psychophysics*, 73, 996-1002.
- Cosman, J. D., & Vecera, S. P. (2012). Context-dependent control over attentional capture. *Journal of Experimental Psychology: Human Perception and Performance*.

ePub ahead of print, October 1, 2012.

Cosman, J. D., & Vecera, S. P. (in press). Learned control over distraction is disrupted in amnesia. *Psychological Science*.

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

Diana, R.A., Yonelinas, A.P., & Ranganath, C. (2007). Imaging recollection and familiarity in the medial temporal lobe: A three-component model. *Trends in Cognitive Science, 11*, 379-86.

Downing, P. E. (2000). Interactions between visual working memory and selective attention. *Psychological Science, 11*, 467-73.

Fiser, J. & Aslin, R.N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science, 12*, 499-504.

Fiser, J. & Aslin, R.N. (2002). Statistical learning of higher-order temporal structure from visual shape-sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 458-67.

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*, 1030-44.

Folk, C. L., Leber, A. B., & Egeth, H. E. (2002). Made you blink! Contingent attentional capture produces a spatial blink. *Perception & Psychophysics, 64*, 741-53.

Folk, C., & Remington, R. (2008). Bottom-up priming of top-down attentional control settings. *Visual Cognition, 16*, 215–31.

- Geng, J. J., & Behrmann, M. (2005). Spatial probability as an attentional cue in visual search. *Perception & Psychophysics*, *67*, 1252–68.
- Hutchinson, J.B., & Turk-Browne, N.B. (2012). Memory-guided attention: control from multiple memory systems. *Trends in Cognitive Science*, *16*, 576-579.
- Kawahara, J.-I. (2010). Identifying a “default” visual search mode with operant conditioning. *Acta Psychologica*, *135*, 38–49.
- Konkel, A., & Cohen, N.J. (2009). Relational memory and the hippocampus: Representations and methods. *Frontiers in Neuroscience*, doi: 10:3389/neuro.01.023.2009.
- Leber, A. B., & Egeth, H. E. (2006). It’s under control: top-down search strategies can override attentional capture. *Psychonomic Bulletin & Review*, *13*, 132–8.
- Leber, A. B., Kawahara, J.-I., & Gabari, Y. (2009). Long-term abstract learning of attentional set. *Journal of Experimental Psychology. Human Perception and Performance*, *35*, 1385–97.
- Mozer, M.C., Shettel, M., Vecera, S.P. (2006). Top-down control of visual attention: A rational account. In Y. Weiss, B. Schoelkopf, & J. Platt (Eds.), *Neural Information Processing Systems 18*, (pp. 923-930). Cambridge, MA: MIT Press.
- Olivers, C. N. L., Meijer, F., & Theeuwes, J. (2006). Feature-based memory-driven attentional capture: visual working memory content affects visual attention. *Journal of Experimental Psychology. Human Perception and Performance*, *32*, 1243-65.
- Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: one phenomenon, two approaches. *Trends in Cognitive Sciences*, *10*, 233-238.
- Ryan, J. D., Althoff, R. R., Whitlow, S., & Cohen, N. J. (2000). Amnesia is a Deficit in

Relational Memory. *Psychological Science*, *11*, 454-461.

Soto, D., Heinke, D., Humphreys, G. W., & Blanco, M. J. (2005). Early, involuntary top-down guidance of attention from working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *31*, 248-61.

Turk-Browne, N. B., Jungé, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General*, *134*, 552-564.

Turk-Browne, N. B., Isola, P. J., Scholl, B. J., & Treat, T. a. (2008). Multidimensional visual statistical learning. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *34*, 399–407.

Vatterott, D. B., & Vecera, S. P. (2012). Experience-dependent attentional tuning of distractor rejection. *Psychonomic Bulletin & Review*, *19*, 871-78.

Figure Captions

Figure 1. An example of the displays and timing parameters used during training and testing sessions, depicting an 'invalid' cue trial (the location of the cue is opposite that of the eventual target, in this case the letter 'H'). On each trial, the cue either matched the target or not.

Figure 2. Data from the training session. Overall reaction times and error rates (base of each column) as a function of whether the cue matched the color that was more (80% of trials) or less (20% of trials) likely to define the target. Error bars represent 95% confidence intervals calculated for each within-subject factor (Cousineau, 2005; Morey, 2008).

Figure 2. Data from the testing session. Overall reaction times and error rates (base of each column) as a function of whether the cue matched the color that was more (80% of trials) or less (20% of trials) likely to define the target during the training session. Error bars represent 95% confidence intervals calculated for each within-subject factor (Cousineau, 2005; Morey, 2008).

Figure 1

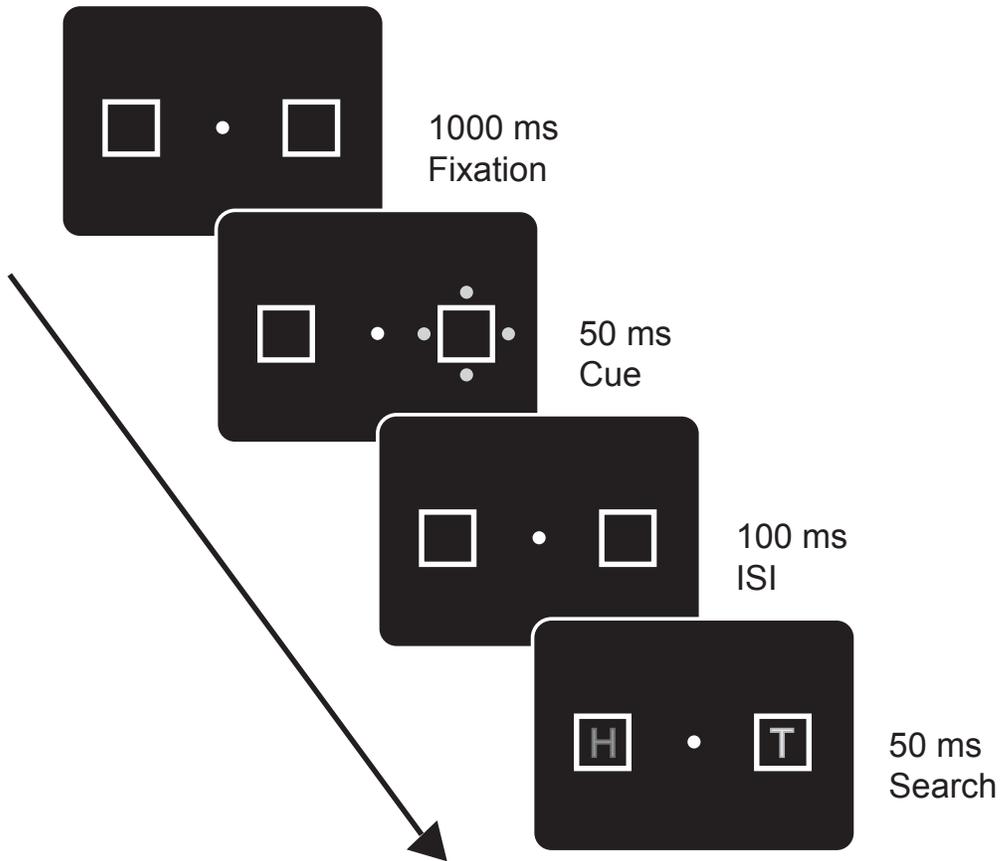


Figure 2

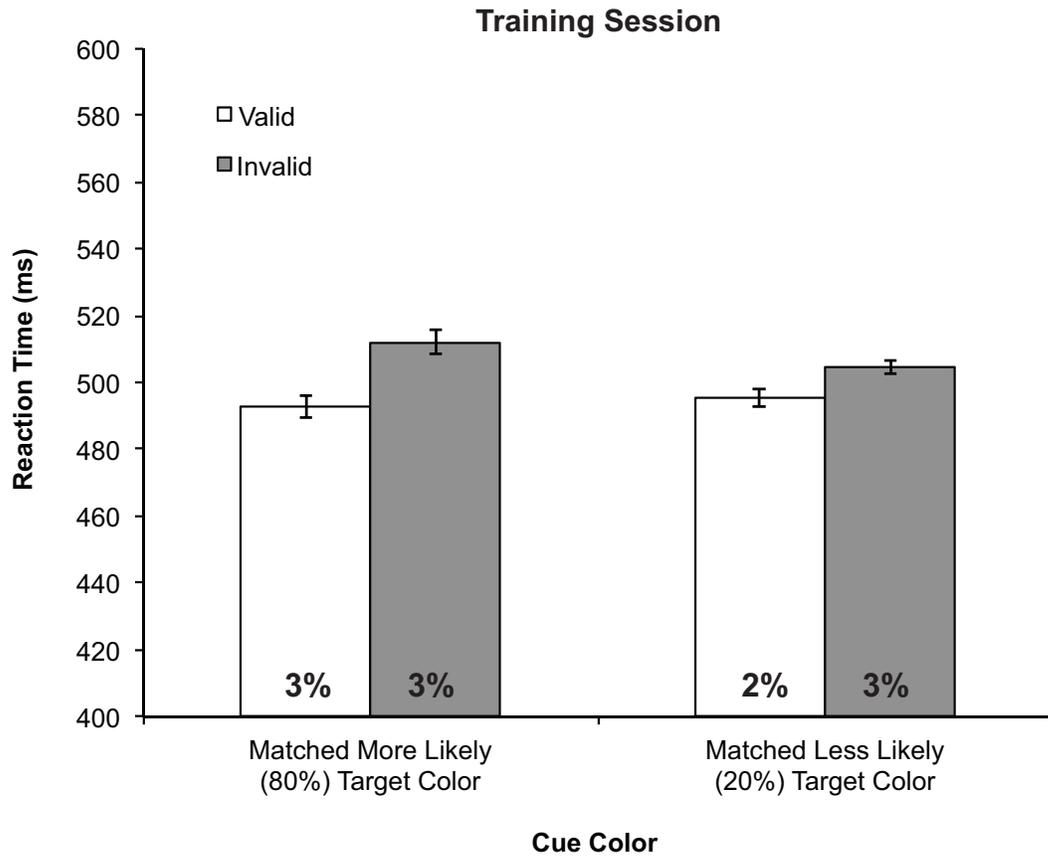


Figure 3

