

Implicit learning of gaze-contingent events

Tom Beesley · Daniel Pearson · Mike Le Pelley

© Psychonomic Society, Inc. 2014

Abstract In a novel implicit learning task, participants responded to a target stimulus that could appear in one of three locations. Unknown to participants, the location in which the target appeared was probabilistically determined on the basis of the location of eye-gaze immediately prior to the appearance of the target. Participants' response times to the appearance of the target in a high-probability location were faster than when it appeared in a low-probability location, revealing that participants were able to learn these gaze-contingent events. Furthermore, there was no difference in the cuing score between those participants classified as aware or unaware of the contingencies on a subsequent forced-choice recognition task. These data suggest the task involves implicit learning of instrumental (action-outcome) contingencies, which has potential implications for our understanding of gaze-contingent processes in social interaction.

Keywords Implicit learning and memory · Eye movements and visual attention

Implicit learning is often defined as learning that occurs incidentally without conscious awareness of the acquired knowledge. In a typical task, participants are given

instructions that are sufficient to complete the task accurately, but make no mention of the task contingencies. For example, in the serial reaction time task (e.g., Cleeremans & McClelland, 1991) participants are told to respond accurately and quickly to a target appearing in one of four locations. Unknown to participants, the target moves in a sequenced manner, such that the next target location can be predicted from the preceding target locations. Despite the incidental nature of the task, participants readily learn such a sequence: response times (RTs) on sequenced transitions are faster than RTs on unsequenced transitions. It has been argued that this learning occurs even though participants show no conscious awareness of the sequence (Cleeremans & McClelland, 1991).

While there is no question that humans can learn in a controlled, conscious manner, the notion that humans also possess a less reflective, unconscious learning system has remained a topic of considerable debate (see Shanks, 2005; 2010). Nevertheless, research has continued to reveal evidence for an implicit learning system. For example, it has been shown that recognition of sequences can be dissociated from the priming effect they generate (Shanks & Perruchet, 2002); that learning of probabilistic contingencies over extended training may occur without awareness (Jiménez, Méndez, & Cleeremans, 1996); and that participants may be unable to control the expression of implicitly acquired (and hence automatically retrieved) knowledge (Destrebecqz & Cleeremans, 2001).

One characteristic of all implicit learning tasks to date is that the to-be-learned contingencies are embedded within the stimuli of the experimental task. For example, in sequence learning, participants perceive and respond to successive movements of a target. It is perhaps unsurprising that in the sparse task-environments used

Electronic supplementary material The online version of this article (doi:10.3758/s13423-014-0720-4) contains supplementary material, which is available to authorized users.

T. Beesley (✉) · D. Pearson · M. Le Pelley
School of Psychology, University of New South Wales, Sydney,
NSW 2052, Australia
e-mail: t.beesley@unsw.edu.au

in traditional implicit learning tasks, participants may notice the contingencies between stimuli. As a result, researchers have attempted to minimize the role of strategic processes, for example by using probabilistic contingencies (e.g., Jiménez et al., 1996), or concurrent tasks to increase working memory load (e.g., Shanks, Rowland, & Ranger, 2005). However, such procedural alterations introduce new difficulties in making valid assessments of implicit and explicit knowledge. Increasing the complexity of the contingencies makes it difficult to understand exactly what the participant has learnt in the task (i.e., the information criterion, Shanks & St. John, 1994), while adding a concurrent task may affect not just learning, but the expression of knowledge (e.g., Frensch, Lin, & Buchner, 1998).

The current experiment uses a novel implicit learning task in which there are no overt contingencies in the presented stimuli, but rather the to-be-learned component is driven entirely by the participant's own behavioral responses. In order to achieve this, the task uses a gaze-contingent design. On each trial, participants are required to respond quickly and accurately to a target appearing in one of three positions. The task is arranged such that participants commonly make a choice to direct their attention towards one of these potential target locations. Crucially, the actual target location is determined by analyzing the gaze location immediately prior to target onset. For example, if the participant was looking at (anticipating a target at) location 1, the target would be most likely to appear in location 2 and would be least likely to appear in location 3. Thus, we assessed whether participants could learn these contingencies between gaze and target position, and to what extent this knowledge was acquired in the absence of awareness.

This novel procedure offers key benefits in the examination of implicit learning. It provides a far simpler set of contingencies than those used in previous tasks, such as artificial grammar learning, sequence learning and contextual cuing (see Shanks 2005; 2010). Awareness of the first-order relationships between gaze and target position can be assessed using a straightforward explicit knowledge test. Gaze-contingent relationships are also, arguably, a non-salient feature of the task; we hypothesize that many participants would not anticipate that their own gaze could act as a cause of task events. In addition, an effect of faster responding to more probable locations would reflect instrumental learning of a relationship between an internally generated action (an eye-movement) and an outcome (the subsequent appearance of the target in a particular location), since there are no external cues that could be

used to predict the upcoming position of the target (that is, the stimuli presented immediately prior to the target appearing are identical on each trial). To our knowledge a purely instrumental implicit learning procedure has not been examined to date. Thus, the task offers a simple means to test the incidental learning of easily defined, yet non-salient contingencies between instrumental actions (gaze preferences) and task events.

Method

Participants Forty-four participants from the University of New South Wales participated for course credit or payment of \$8. Participants received bonus payment (see below) totaling on average \$6.03.

Apparatus and stimuli A Tobii TX300 Eye Tracker (Tobii Technology, Sweden) recorded gaze at a sampling rate of 300 Hz, with spatial resolution of 0.15°. Participants used a chinrest ~60 cm from the 23-inch monitor (58.4 cm; resolution of 1920×1080; refresh rate of 60 Hz).

The experiment was programmed using the Psychophysics Toolbox (<http://psychtoolbox.org>) for MATLAB. Figure 1A shows a screenshot following a correct response and presentation of feedback. The background was black and all stimuli were white. Three circular rings of diameter ~2.0 degrees of visual angle (hereafter °) were arranged in a triangle, with the centre of each circle at a vertical distance of ~8.7° from the centre. The left and right circles were offset from the horizontal midline by ~15.1°; the top circle was presented on the horizontal midline. The target was a “T” rotated by 90° or 270°, determined at random on each trial. Auditory feedback was presented through headphones and consisted of brief (~500 ms) sound clips of a cash register opening (large bonus), a coin hitting a table (small bonus), or a buzzer (incorrect/timeout).

Design Trials commenced with a pre-target phase in which the three circles marking the stimulus locations were presented for 1000 ms. During this pre-target period, a *stimulus-gaze event* was registered whenever gaze fell within an invisible square region (subtending ~6.3°) centered on the locations, for a continuous period of 50 ms. The subsequent target position was determined on the basis of the final stimulus-gaze event.

Table 1 shows the probability of the target appearing in each of the three locations, as a function of the location of the final stimulus-gaze event. There were three possible types of trial. On *match trials* (probability = .3), the target appeared in the location of the stimulus-gaze

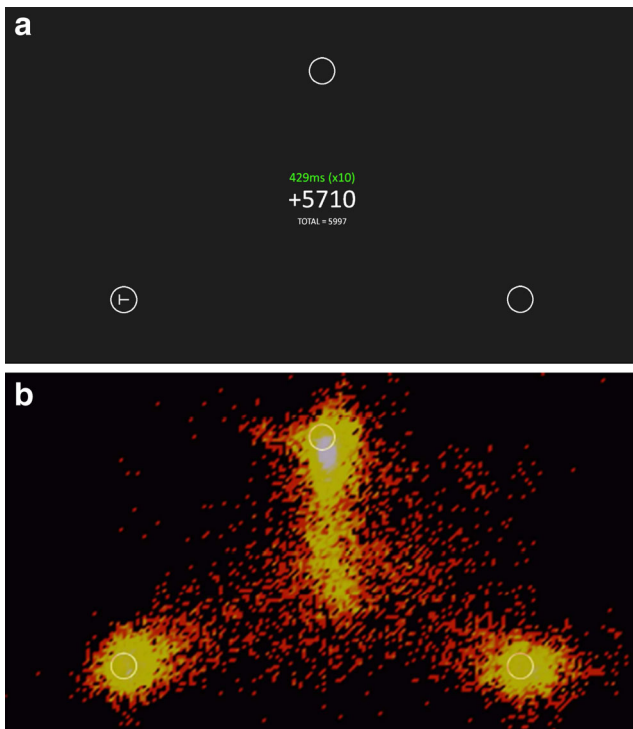


Fig. 1 (A) A screenshot of the main experimental task following a correct response to the target stimulus. Feedback for the performance incentive is shown in the centre of the screen. (B) The distribution of final fixations during the pre-target period, accumulated for all participants. More densely fixated regions are represented by lighter colors

event. On mismatch trials, the target appeared in one of the other two locations. Critically, the target was more likely to occur in one of these unattended locations (*frequent-mismatch trials* [probability = .6]) than the other (*infrequent-mismatch trials* [probability = .1]).

Table 1 Contingencies between the Final Stimulus-Gaze Event During the Pre-Target Period and the Subsequent Target Location

	Target appears in 1	Target appears in 2	Target appears in 3
Gaze on 1	.3	.6	.1
Gaze on 2	.1	.3	.6
Gaze on 3	.6	.1	.3

Note: “Gaze” reflects the stimulus location (1, 2 or 3) in which the final stimulus-gaze event was registered (see main text). “Match trials” are those in which the target appears in the same location as the final stimulus-gaze event was registered (occurring with a probability of .3). Mismatch trials are those in which the target appears in a different location from the final stimulus-gaze event: “frequent mismatch” trials are those occurring with a probability of .6; “infrequent mismatch” trials are those occurring with a probability of .1

Each position acted as the target location on frequent-mismatch trials when the final stimulus-gaze event was registered in one location, and on infrequent-mismatch trials when it was registered in the remaining location.

In summary, if the final stimulus-gaze event was on location 1, the target was most likely to appear in location 2; if the final stimulus-gaze event was on location 2, the target was most likely to appear in location 3; and if the final stimulus-gaze event was on location 3, the target was most likely to appear in location 1. For all participants, location 1 was the circle at the top of the screen. For half of the participants, location 2 was the circle on the right, and location 3 was the circle on the left. For these participants the frequent-mismatch contingencies were arranged in a clockwise rotation. For the remaining participants, location 2 was on the left, and location 3 was on the right; this was the counter-clockwise arrangement.

Trial types were determined independently for each location, such that for every 10 trials in which the final stimulus-gaze event fell on any one location, the target would occur in the match location 3 times, the frequent-mismatch location 6 times, and the infrequent-mismatch location once. If no stimulus-gaze events were registered during the pre-target phase, target location was randomly determined (and this did not count as one of the 10 trials described in the preceding sentence).

Target responses were made using the N (for targets rotated 90°) and C keys (270°) of a standard keyboard.

Procedure Participants completed a five-point eye-tracker calibration. Instructions detailed the task, including an example display showing the correct response for each target orientation. Participants were informed that they would earn points during the task, and that a bonus of \$1 was given for every 100,000 points attained. In order to encourage fixations on one of the locations during the pre-target period, the instructions outlined that the best strategy for maximizing the number of 10x bonus outcomes (see below) was to try to anticipate where the target was going to occur by attending to one of the locations prior to the target appearing.

The experiment consisted of 400 trials. Trials commenced with a pre-target phase in which the three circles marking the stimulus locations were presented for 1000 ms. The target then appeared in one location. Following a response, or a period of 2000 ms (timeout), visual feedback (see Fig. 1A) was provided for 2000 ms and any auditory feedback was delivered. Visual feedback consisted of RT, the bonus multiplier (see below), the number of points awarded for the trial, and the total points accumulated. The color of the text displaying the

RT and bonus multiplier varied from dull to bright green according to the bonus level. For RTs greater than 1000 ms, no points were awarded and the visual feedback showed “(no points – too slow)”. For RTs less than 1000 ms, participants received points of $(1000 - RT) \times \text{bonus}$. For RTs less than 700 ms a 2x bonus was applied; for RTs less than 450 ms a 10x bonus was applied. Pilot work revealed that this threshold for the 10x bonus increased the likelihood of participants adopting the strategy of selectively attending one of the stimulus locations to maximize reward. For inaccurate responses or timeouts, “ERROR!” or “TIMEOUT! RESPOND FASTER!”, respectively, was displayed in red in the center of the screen and participants lost 5000 points. The points total could never fall below zero. The inter-trial interval was 500 ms.

A 20 second rest-break was given every 50 trials, during which a summary was displayed of the bonuses attained, and the corresponding points won and penalty points lost. Participants were reminded of the threshold to achieve the 10x bonus, and that the best strategy was to attend to a location during the pre-target period. Trials re-started automatically after breaks.

Immediately after the main task, participants were given an awareness questionnaire (see [Appendix](#)). The questionnaire outlined the gaze-contingent nature of the task, and gave participants the choice of two possible relationships between the position of their eye gaze and the frequency of target location (i.e., the clockwise and counter-clockwise arrangements of the frequent-mismatch contingencies; see Design). Participants gave a confidence rating for their decision, from 1 (“I was guessing”) to 10 (“I’m certain”).

Results

During the pre-target period participants made, on average, 1.4 fixations (standard error of the mean [SEM] = .11); the average duration of fixations was 263 ms (SEM = 6 ms). Figure 1B shows a density plot of the final fixations during the pre-target period, revealing that participants favored the top stimulus location.¹ On average, a stimulus-gaze event was registered on 73 % of

trials (SEM = 4 %). The mean lag between the final stimulus-gaze event and target onset was 76 ms (SEM = 12 ms). Four participants had fewer than 25 % of trials in which this lag was less than 500 ms (the mean for all participants was 71 %, SEM = 5 %). Since these four participants rarely looked at stimulus locations during the pre-target period they were excluded. The mean lag after the removal of these participants was 57 ms (SEM = 8 ms) and on 61 % (SEM = 4 %) of trials participants’ gaze was on a stimulus location at the moment the target appeared. Trials with incorrect responses (8 %), trials with no detected stimulus-gaze event (20 %), and trials with a lag greater than 500 ms (2 %) were removed from the analysis of RTs.²

Mean RTs were faster on match trials ($M = 467$ ms, $SEM = 11$) than mismatch trials ($M = 637$ ms, $SEM = 10$ ms), $t(39) = 13.51$, $d = 2.14$, $p < .001$. This is unsurprising, since on match trials participants were recently looking at the location in which the target appeared (and for the majority of trials at the precise moment it appeared). More importantly, responses were faster on frequent-mismatch trials ($M = 619$ ms, $SEM = 10$ ms) than infrequent-mismatch trials ($M = 654$ ms, $SEM = 11$ ms). An ANOVA with factors of frequency (frequent-mismatch vs. infrequent mismatch) and epoch (of 100 trials) revealed a main effect of frequency, $F(1,39) = 21.72$, $\eta_p^2 = .36$, $p < .001$, suggesting that when the target appeared in a mismatch location, responses were faster when the target occurred in the more predictable location. This suggests that participants learnt the contingencies between pre-target gaze and target location. A significant interaction effect revealed that the effect of frequency emerged with training, $F(3,117) = 3.02$, $\eta_p^2 = .07$, $p = .032$; the effect was not present in epoch 1, $t(39) = 1.65$, $d = .26$, $p = .11$, but was present in epochs 2-4, $t_s(39) \geq 3.19$, $d_s \geq .50$, $p_s \leq .003$. The main effect of epoch was not significant, $F < 1$.

Individual awareness scores were calculated from participants’ confidence rating. If the correct arrangement of frequent-mismatch contingencies was selected, the awareness score was equal to the confidence score.

¹ See [supplemental material](#) for detailed analysis of the stimulus-gaze events.

² The threshold for stimulus-gaze event lags of 500 ms ensures that our analysis examines only those trials in which there was reasonable contiguity between the stimulus-gaze event and the target presentation. The duration of 500 ms also ensures that the analysis retains a large proportion of the trials. However, it should be noted that the critical effect of faster responding on frequent- over infrequent-mismatch trials is observed across a range of lags, including when the analysis is restricted to trials in which there was no lag (0 ms), $t(39) = 4.58$, $d = .84$, $p < .001$.

However, if the incorrect arrangement was selected, the awareness score was the confidence rating multiplied by -1 . This gave an awareness measure ranging between -10 and 10 , where a negative score indicates the participant was unaware of the relationship between their attention and the occurrence of the target in the frequent-mismatch locations.

Overall, 27 out of 40 participants produced a positive awareness score, which a binomial test revealed was greater than chance, $p = .040$. These participants were classified as “aware”, and the remaining 13 participants (with negative scores) as “unaware”. Figure 2A plots the RTs for the different trial types as a function of classification on the awareness test. The pattern is very similar in the two sub-groups. Notably, RTs were significantly faster

on frequent-mismatch trials compared to infrequent-mismatch trials for both aware, $t(26) = 4.03$, $d = 1.36$, $p < .001$, and unaware participants, $t(12) = 2.65$, $d = .73$, $p = .021$. Our measure of gaze-contingent learning is given by the difference in RT between infrequent-mismatch and frequent-mismatch trials. This *mismatch cuing score* (aware = 32 ms; unaware = 40 ms) did not differ between the two groups, $t < 1$. Figure 2B shows a scatterplot of awareness scores against mismatch cuing score; there was no correlation between these measures, $r(40) = -.057$, $p = .73$. These data therefore suggest that learning was unrelated to conscious awareness of the gaze-contingent relationships.

Discussion

In a novel implicit learning task, participants were trained to anticipate an upcoming target, the position of which was contingent upon the location of the participant’s gaze prior to it appearing. Data from the vast majority of participants were consistent with having learnt this relationship: Figure 2B shows that 34 out of 40 participants had a positive mismatch cuing score (an RT benefit for the more probable target locations).

The contingent relationships in this task were driven purely by participants’ own behavior during the pre-target period. Consequently, such learning must rely entirely on instrumental associations (action-outcome learning), since there were no valid external (Pavlovian) cues that could signal the target location; to our knowledge, this is the first time instrumental contingency learning, independent of external cues, has been demonstrated in an implicit learning task.

There are at least two possible accounts of what might be learnt in this task. On the basis that attention correlates closely with eye movements in tasks with spatially discrete cues (Deubel & Schneider, 1996), these data may suggest that participants are able to associate their attentional localization with an upcoming contingent event. The implication of this account is that attentional movements can themselves act as internal cues, generating a series of expectancies for events within the environment. If this account is true, it has important implications for our understanding of the role of attention in the cognitive system, since it suggests that attention is not confined merely to the role of filtering incoming information, but is subject to the laws of instrumental conditioning like any other behavioral response (see also Le Pelley, Mitchell &

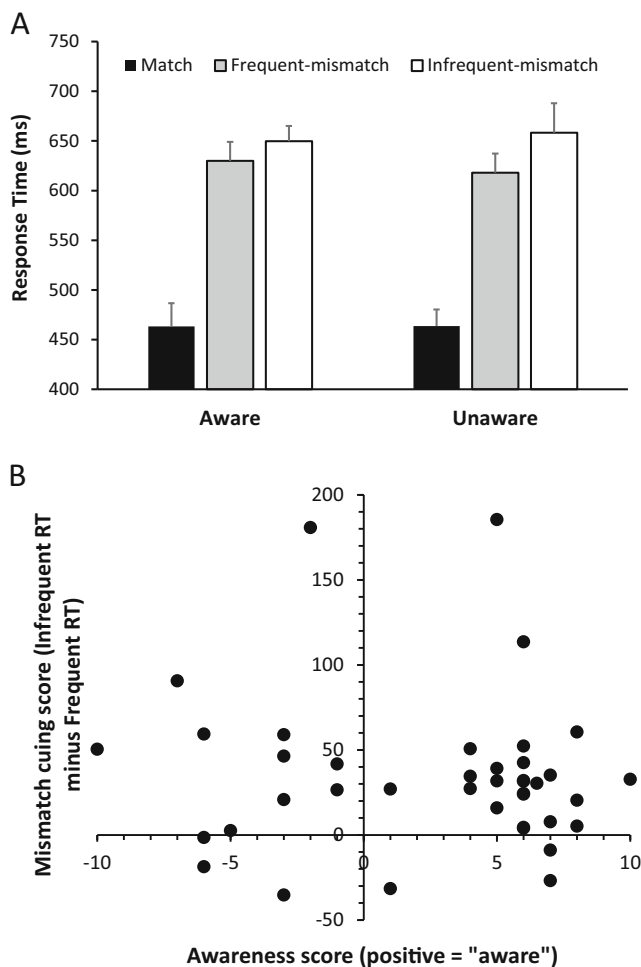


Fig. 2 (A) Mean response times (RTs) to targets across the three trial-types for participants classified as aware and unaware. (B) Scatterplot of the “mismatch cuing effect” (mean RT on infrequent mismatch trials minus mean RT on frequent mismatch trials) as a function of participants’ awareness score

Johnson, 2013). Alternatively, participants may acquire a preference for making eye movements in line with the clockwise or counter-clockwise rotation of the probabilistic contingencies in the task, depending on the condition with which they are trained. Provided these movements are timed to coincide with the occurrence of the target, they will be rewarded with faster response times in the task (and a slightly larger incentive payout). Differentially rewarding such eye-movements would lead to an increase in vigour for those eye-movements which match the trained contingencies.³ The current data do not allow us to decide between these two accounts.

Following the task we examined whether participants were aware of the contingencies between gaze and target position, using a simple forced-choice recognition test. The fact that more than half of the participants selected the correct set of contingencies suggests that some participants were aware of these relationships. However, there was no evidence of a correlation between the size of the mismatch cuing effect and awareness scores. Furthermore, those “unaware” participants that selected the incorrect set of contingencies in the awareness test, showed a significant mismatch cuing effect, equivalent to that observed in “aware” participants.

Clearly some participants became aware of the contingencies in this task, and in this respect our data are compatible with demonstrations of above-chance recognition in other implicit learning tasks (e.g., Shanks & Johnstone, 1999). However, the comparable cuing effects in “aware” and “unaware” participants, and the lack of a correlation between the measures, suggests that awareness was not a necessary condition for learning in the task. Perhaps it is the case then that conscious expression of knowledge may emerge gradually; it may be necessary for the gaze-target associations to strengthen beyond some threshold, after which the knowledge can be verbally expressed, with this threshold not reached in the “unaware” participants. While this view of awareness as an emergent property of learning has received support in the implicit learning literature (e.g., Cleeremans & Jiménez, 2002), it would

be remiss not to consider an alternative argument, namely that the two measures of learning (cued RT and recognition) differ in their sensitivity to detect the underlying knowledge. While our measure of cuing is derived from many RTs, our measure of awareness is derived from a single reflective judgment. The effects of a noisy memory retrieval process may well be somewhat greater in the latter case, and may lead to such a dissociation, even when the measures tap a single knowledge base (Shanks & Perruchet, 2002). Future work will attempt to maximise the sensitivity of the awareness measure.

The current task has examined gaze-contingent cuing of a target location, yet exactly how such learning leads to a performance benefit is unknown. One possibility is that the attentional preference primes subsequent shifts of attention towards the contingent location. Alternatively, it may result in the enhanced detection of the target at an anticipated location. Similar debates within the contextual cuing literature have uncovered important details about these performance processes (e.g., Kunar, Flusberg, Horowitz, & Wolfe, 2007) and similar explorations are likely to be important in understanding gaze-contingent cuing. Furthermore, it will be worth examining whether internal attentional preferences can serve as cues for other behavioral responses; the allocation of attention to particular regions may drive expectations for appropriate actions to upcoming events.

An understanding of gaze-contingent cuing may also be of importance in understanding social interaction. During face-to-face communication, the gaze direction of a respondent has an effect on their perceived trustworthiness (Mason, Tatkow, & Macrae, 2005), and in young infants, changes in behavior (e.g., smiling) are observed when gaze is diverted away from the eyes (Symons, Hains, & Muir, 1998). In these social settings gaze movements lead to observable events in the world (behavioral changes in other humans). The implication is that gaze-contingent cuing may be driven by an automatic learning process that contributes to social interaction. It may be the case that these fundamental learning processes are impaired in populations thought to have deficits in social interaction, such as those who score highly on the Autism-Spectrum Quotient.

³ We thank Ian McLaren for suggesting this alternative account.

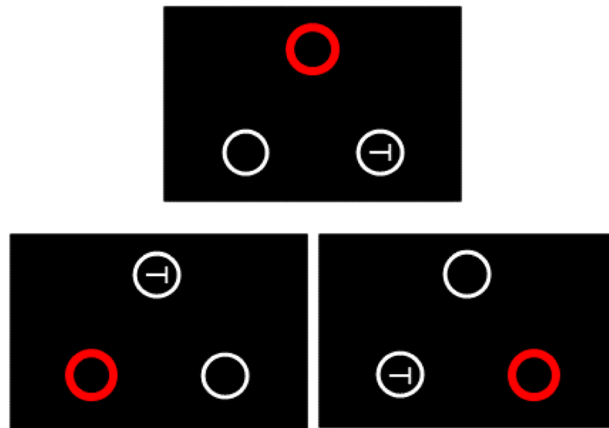
Appendix

Forced-choice recognition test for awareness of task contingencies

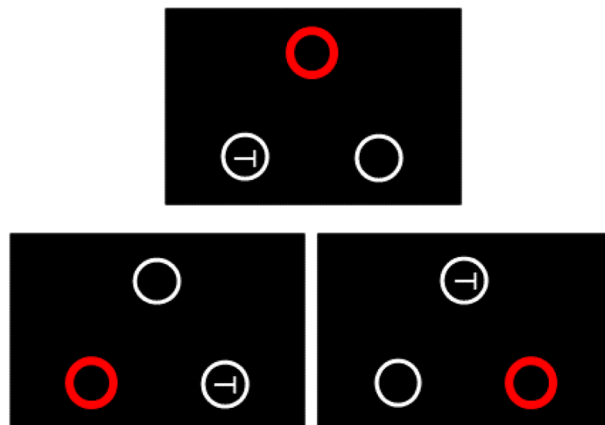
In this task, the position that the target appeared in was determined by where you were looking immediately prior to that target appearing. Below we have presented two possible options of how this task operated. Consider each one in turn and select one option. If you are unsure of what you have to do, please ask the experimenter.

Option 1

When I was looking at any of the three locations in red, the target would usually appear in the next clockwise location.

**Option 2**

When I was looking at any of the three locations in red, the target would usually appear in the next anti-clockwise location.



Please tick the option you think reflects the task you just completed. How confident are you in your decision (1 "I was guessing" to 10 "I'm certain")? Answer: _____

References

- Cleeremans, A., & Jiménez, L. (2002). Implicit learning and consciousness: A graded, dynamic perspective. In R. M. French & A. Cleeremans (Eds.), *Implicit Learning and Consciousness* (pp. 1–40). Hove, UK: Psychology Press.
- Cleeremans, A., & McClelland, J. L. (1991). Learning the structure of event sequences. *Journal of Experimental Psychology: General*, *120*, 235–253.
- Destrebecqz, A., & Cleeremans, A. (2001). Can sequence learning be implicit? New evidence with the process dissociation procedure. *Psychonomic Bulletin & Review*, *8*, 343–350.
- Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research*, *36*, 1827–1837.
- Frensch, P. A., Lin, J., & Buchner, A. (1998). Learning versus behavioral expression of the learned: The effects of a secondary tone-counting task on implicit learning in the serial reaction task. *Psychological Research*, *61*, 83–98.
- Jiménez, L., Méndez, C., & Cleeremans, A. (1996). Comparing direct and indirect measures of sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 948–969.
- Kunar, M. A., Flusberg, S. J., Horowitz, T. S., & Wolfe, J. M. (2007). Does contextual cuing guide the deployment of attention? *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 816–828.
- Le Pelley, M. E., Mitchell, C. J., & Johnson, A. M. (2013). Outcome value influences attentional biases in human associative learning: Dissociable effects of training and of instruction. *Journal of Experimental Psychology: Animal Behavior Processes*, *39*, 39–55.
- Mason, M. F., Tatkov, E. P., & Macrae, C. N. (2005). The look of love: Gaze shifts and person perception. *Psychological Science*, *16*, 236–239.
- Shanks, D. R. (2005). Implicit learning. In K. Lamberts & R. L. Goldstone (Eds.), *Handbook of Cognition* (pp. 202–220). London: Sage.
- Shanks, D. R. (2010). Learning: from association to cognition. *Annual Review of Psychology*, *61*, 273–301.
- Shanks, D. R., & Johnstone, T. (1999). Evaluating the relationship between explicit and implicit knowledge in a sequential reaction time task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 1435.
- Shanks, D. R., & Perruchet, P. (2002). Dissociation between priming and recognition in the expression of sequential knowledge. *Psychonomic Bulletin & Review*, *9*, 362–367.
- Shanks, D. R., Rowland, L. A., & Ranger, M. S. (2005). Attentional load and implicit sequence learning. *Psychological Research*, *69*, 369–382.
- Shanks, D. R., & St. John, M. F. (1994). Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences*, *17*, 367–395.
- Symons, L. A., Hains, S. M. J., & Muir, D. W. (1998). Look at me: Five-month-old infants' sensitivity to very small deviations in eye-gaze during social interactions. *Infant Behavior and Development*, *21*, 531–536.