Working Memory and IQ Predict Plan-based Control: Evidence from Novel Serial Reaction

Time and Reinforcement Learning Paradigms

Kuipers, M.[1]

De Kleijn, R. E. [1]

1. Cognitive Psychology Unit, Institute of Psychology, Leiden University, Leiden, The Netherlands

Abstract

Most daily human behaviors can be described in terms of sequential actions. Although this issue has been extensively examined in psychological and neuroscientific research, much remains unknown about how humans learn and produce these actions. A paradigm prominent in the study of sequence learning is the serial reaction time (SRT) task, which aims to explain complex skill acquisition. In the present study we shed light on the mechanisms underlying sequence learning by examining determinants of control mode in a cued SRT task and performance in a reinforcement learning task in which plan-based control is paramount. It was hypothesized that individual differences in visuospatial working memory capacity, IQ, locus of control, and personal need for structure, are predictive of sequence learning performance and action control mode. Our results indicate that sequence learning is not dependent on personality characteristics, but rather on cognitive capabilities. We further reproduced the Nissen and Bullemer (1987) speedup in movement times, but were unable to replicate the frequency effects observed in Tubau, Hommel, and López-Moliner (2007). Although experimental limitations are acknowledged we consider our study to be a valuable contribution to the current discussion and encourage further research into the mechanisms underlying human sequence learning.

Keywords: Action control, Action plan, Frequency effects, Reinforcement learning, Sequence learning, Serial reaction time task, Sequential action, Trajectory

Working Memory and IQ Predict Plan-based Control: Evidence from Novel Serial Reaction

Time and Reinforcement Learning Paradigms

How humans learn sequences has been a long-standing research problem in psychology

and is currently a major topic in neuroscience. A wide variety of studies have aimed answering

this question with topics ranging from implicit sequence learning (Cleeremans & McClelland,

1991; Destrebecqz & Cleeremans, 2001; Jiménez & Méndez, 1999; Meulemans, Van der Linden,

& Perruchet, 1998; Nissen & Bullemer, 1987) and explicit sequence learning (Cohen, Ivry, &

Keele, 1990; Rauch et al., 1995; Schendan, Searl, Melrose, & Stern, 2003), to artificial grammar

learning (Dienes, Broadbent, & Berry, 1991; Knowlton & Squire, 1994, 1996). Throughout our

lives we are surrounded by behavioral sequences: from learning to walk and performing the

tango, to adding numbers and solving linear equations. Performing such sequences can be

demanding at first, but with enough practice they can be performed almost effortlessly. In this

context a sequence is considered to be a set of related events, movements, or items that follow

each other in a particular order.

In the early stages of questioning how people acquire behavioral sequences James (1890)

theorized that actions can be 'chained' by sequentially perceiving the sensory feedback of the

previous action. With experience, each action's sensory effect (e.g. the feeling of placing your

laptop on a desk) becomes associated with the next action component (e.g. opening the lid of

your laptop) through stimulus-response learning, until the sequence comes to an end. Put

differently, each subsequent action was thought to be automatically triggered by, for example,

response-produced afferent information from the muscles of the previous action. An external

stimulus would thus be sufficient to trigger the further performance of the sequence without

much need for conscious control. As explained in James (1890, p. 359), "… if such a reaction

has many times occurred we learn what to expect of ourselves, and can then foresee our conduct, even though it remain as involuntary and uncontrollable as it was before". Sequence learning could thus be interpreted as experience-dependent improvement of our sensory system to respond to stimuli.

In spite of starting consensus that sequential learning calls upon response related information, Hazeltine (2002) found that a change in the response sequences does not hinder performance if the environmental consequences of new responses remain identical to those when the sequence was still being learned. This has been cited as evidence that sequence learning could be neither stimulus-based nor response-based, but rather be based on the formation of a plan. In accordance, Münsterberg (1892) posited that the associative account in James' response-chaining hypothesis (James, 1890) is inadequate to construe sequential action because a directional component is needed to effectively perform actions in the correct order. He argued that the learning process of action sequences relies on the acquisition of a motor program, but failed to explain why the directional component does not apply in motor learning. Still, by pioneering the idea of a cognitive structure that governs execution of sequential actions he provided a provocative theoretical alternative to the chaining theory. By converging James' and Münsterberg's approaches toward sequence learning, Tubau, Hommel, and López-Moliner (2007) argued that one of these two approaches does not need to be right or wrong. Rather it was suggested that James' stimulus-driven approach of sequence learning and Münsterberg's program hypothesis reflect two alternative modes of executive control: stimulus-based control and plan-based control, respectively. Several lines of diverse research have not only provided imperative evidence that executive control in sequence learning can indeed be identified as either stimulus-based or plan-based (Herwig, Prinz, & Waszak, 2007), but also that a shift from

stimulus-based to plan-based control develops as sequence learning progresses (Hoffmann &

Koch, 1997). Under stimulus-based control a large amount of cognitive control is delegated to

external stimuli, which is similar to James' chaining theory positing that the execution of an

action triggers a following action. Stimulus-based control is characterized by the automatized

response to external stimuli, reflected by absence of developing explicit sequence knowledge

(Tubau et al., 2007). As a consequence, not much of the sequence is learned during stimulus-

based control. Contrarily, the plan-based control mode is thought to rely on the construction of

an action plan (Hommel, 2003; Luria, 1961; Miller, Galanter, & Pribram, 1960), which implies

that plan-based representations are generated internally (Vygotsky, 1986; Zelazo, 1999) and can

be expressed explicitly (Tubau et al., 2007).

　　　Sensitivity to frequency information (i.e. facilitation of responses to frequent compared

to infrequent sequence transitions) has been thought to be informative about the current action

control mode during sequence production. Support for this 'frequency effect' comes from Tubau

et al. (2007) in which the shift between executive control modes was investigated using a serial

reaction time (SRT) paradigm that required participants to learn an underlying sequence pattern.

In this paradigm the letter 'X' appeared to either the right or the left relative to the center of the

screen in accordance to the continuously repeating R-L-R-R-L-L-R-L sequence (where R is right

and L is left). Participants received either incidental instructions in which the experiment was

introduced as one exploring the effect of training on reaction time, or they received intentional

instructions in which the participant was informed about the presence of a repeating sequence of

locations and the goal was to discover the structure of the sequence. By asking participants

whether they had noticed any repeating sequence, and in cases of an affirmative answer asking

them to write down the repeating sequence, Tubau et al. (2007) were able to measure each

participant's knowledge about the sequence. In this sequence (R-L-R-R-L-L-R-L), in which

stimulus alternations occur more often than repetitions, it was found that implicit learners (i.e.

participants without explicit knowledge of the underlying sequence pattern) were notably

affected by the pattern's frequency information. Specifically, it was found that responses to

stimulus alternations were faster than to repetitions. Explicit learners (i.e. participants with

explicit knowledge of the underlying sequence) on the other hand were much faster than implicit

learners and did equally well on both the repetition and alternation responses. Tubau et al.

(2007) speculated that this difference in performance between both types of learners can be

explained by a difference in action control mode. That is, a stimulus-based control mode in

implicit learners, and plan-based control mode in explicit learners which diminishes the impact

of local stimulus-based response bias.

   The majority of studies concerning sequence learning have focused on learning in a cued

paradigm (i.e. participants are presented with a target and instructed to respond to the target

immediately after observing it), similar to the Tubau et al. (2007) study. However, there is

evidence for the notion that sequence learning does not always occur by linking stimulus-

response sets. Instead, it could be argued that behavioral sequence acquisition is better defined

as an exploratory process in which people undertake multiple trials before succeeding. Recently,

the SRT paradigm was modified into a reinforcement learning paradigm (Kachergis, Berends, de

Kleijn, & Hommel, 2016) making it possible to examine how individuals discover adaptive

behaviors in stable environments. Whereas in the original SRT task (cued sequential action

paradigm) participants earn points by simply following the presented cues as fast as possible, its

reinforcement learning adaptation does not provide any cues and requires participants to explore

four alternatives until the correct target was found, receiving feedback (score increment or

reduction) after each response. Participants were instructed to maximize their score, which was displayed continuously at the top of the screen. Noteworthy, participants' final scores revealed a bimodal distribution, with about half of the participants scoring very high, and the other half scoring below average. We argue that this finding could be a reflection of individual differences with regards to personality characteristics, cognitive capabilities, and preferred use of control mode strategy. Even though Kachergis et al. (2016) examined both the trajectory SRT paradigm (Kachergis, Berends, Kleijn, & Hommel, 2014) and its reinforcement learning adaptation, due to the design of their study it was not possible examine whether individual differences could explain results that are common in both tasks.

In the present study we investigate the determinants of action control modes in both a cued sequential action and an exploratory reinforcement learning paradigm. While Tubau et al. (2007)'s SRT task is keypress based, its trajectory adaption (Kachergis et al., 2014) is mouse movement based and allows for continuously tracking the mouse cursor location, thus having the advantage being able to record predictive movements. As the plan-based control mode is characterized by making predictive movements in the absence of cues (Nattkemper & Prinz, 1997) detection of predictive movements (i.e. mouse movements toward the next stimulus in the sequence) can thus be considered indicative of the current action control mode employed. Due to this advantage, the novel trajectory SRT task will be used in the present study to identify action control modes among participants. By complementing our experimental design with the SRT reinforcement learning adaptation from Kachergis et al. (2016) we attempt to reproduce their bimodal score distribution, and followingly examine for the first time whether the distinctive distribution peaks are a product of different action control mode strategies. Furthermore, even though Tubau et al. (2007) have shown that action control modes can be

evoked experimentally by using symbols to help the creation of an action plan, individual

differences could also play a pivotal role in the use of control mode strategy. Earlier research

has shown that performance in a visuospatial working memory task predicts implicit motor

sequence learning (Bo, Jennett, & Seidler, 2011) and that performance in a memory updating

task is predictive of performance in a visuospatial sequence learning task (Martini, Furtner, &

Sachse, 2013). Considering this, we posit that working memory capacity could be a potential

determinant of one's action control mode. Additionally, we argue that fluid intelligence could

play a pivotal role in sequence learning as it is possible that some individuals do not have the

cognitive capabilities to form a long action plan, constraining these individuals to employ a

stimulus-based action control mode strategy. We hypothesize that there is a relation between

these measures of cognitive capabilities and the action control mode employed, where

participants with either a relatively low working memory capacity or fluid intelligence show

signs of a stimulus-based control mode, as reflected by the absence of acquiring explicit

sequence knowledge and scarcely displaying predictive movements. Further, we believe that an

individual's locus of control could also play a pivotal role in the formation of an action plan.

Whereas individuals with a strong internal locus of control tend to believe that events in their

lives are the result of their own actions, individuals with a strong external locus of control tend to

tend to believe that such events are the consequence of external forces beyond their control. We

posit that individuals with an internal locus of control tend to display characteristics of having

engaged in plan-based control, while individuals with an external locus of control are more likely

to exhibit the signs of having engaged in stimulus-based control. Moreover, individual

differences in need for structure could also contribute to action plan formation. While some

individuals could have no need to establish structure in their daily lives, others may have this

need and may prefer to actively predict future situations according to a plan instead of waiting for unforeseen situations or stimuli to arrive (Neuberg & Newsom, 1993). It is expected that participants having a high need for structure actively look for structure in sequential actions, thus showing predictive movements and the ability to verbally report explicit sequence knowledge, while participants with a low need for structure are not. Finally, we hypothesize that sensitivity to frequency information is dependent on the development of explicit sequence knowledge. It is expected that both implicit and explicit sequence learners start with similar levels of sensitivity to frequency information, but that over time only explicit learners develop an action plan which diminishes the impact of local stimulus-based response bias.

Method

Participants

Forty undergraduate students at Leiden University (age mean = 20.79 years, SD = 2.34 years; 13 males and 27 females) participated as part of gaining course credit. All participants spoke English fluently and signed informed consent prior to their inclusion in the study. The research protocol for this study was approved by the Psychology Research Ethics Committee at Leiden University. Participants were required to meet the following inclusion criteria: 1) 18-30 years of age, 2) absence of taking drugs or medication, 3) absence of (history of) psychiatric or neurological disorders. The total duration of the experiment was approximately 80 minutes, depending on task performance.

Materials and measures

All tasks were conducted using E-Prime software version 2.0.10.356 (Psychology Software Tools, Pittsburgh, PA) and presented on a 17" monitor set to 1024 x 768 resolution at a

distance of about 70 cm from the participants' eyes.

Working memory capacity. To assess working memory capacity we used the visuospatial working memory task from Bo et al. (2011), which was a modified version of the task used by Luck and Vogel (1997). The stimuli consisted of 2-8 (array size) colored circles which were presented in varying colors (red, orange, yellow, green, blue, violet, pink, white, black, and brown) on a white background. For each trial, the test array was either the same as the sample array or different with only one of the colors changed. Participants were instructed to indicate whether the test array was the same (response 'S') or different (response 'D') from the sample array by keypress (Figure 1). Therefore, this task relied on the detection of a change in color at different locations. In this task all colored circles were arranged along an invisible circle around a fixation cross. Working memory capacity was computed using the formula: $K = $ Size of the array $*$ (observed hit rate - false alarm rate; Vogel & Machizawa, 2004). The average K over all array sizes was determined to compute each participant's working memory capacity. Participants completed 140 trials in total, during which at the halfway point they received a short break of 1 minute.

Fluid intelligence. To assess fluid intelligence we administered a 10-minute version of the Raven's Standard Progressive Matrices (SPM; Raven, Court, & Raven, 1998). Raven's SPM non-verbally assesses the participant's capacity for analyzing and solving problems, abstract reasoning and the ability to learn. The task contains reasoning items ordered with increasing difficulty, demanding greater cognitive capacity to solve. Given eight geometric figures, the participant is asked to identify the ninth, missing, geometric figure that completes a pattern. All items were presented in black on a white background. Raven's SPM is widely used as a measure of general mental ability and fluid intelligence. IQ scores were estimated by

normalizing the number of correct responses of all participants to a distribution with mean 100 and standard deviation 15. The task has been shown to have excellent internal consistency reliability, convergent validity, and criterion-related validity (Raven, Raven, & Court, 2000).

Personal need for structure. In order to measure the extent participants prefer structuralizing and organizing experiences (without referring to social or political issues) in their daily lives, we administered the Personal Need for Structure questionnaire (Thompson, Naccarato, & Parker, 1989). The scale consists of 12 items (e.g. "I enjoy having a clear and structured mode of life") and assesses people's desire for structure (items 3, 4, 6, and 10) and response to a lack of structure (items 1, 2, 5, 7, 8, 9, 11, 12). The Personal Need for Structure scale is based from the assumption that the human ability to reduce the uncertainty of situations is associated with the ability to manage unknown situations. Participants responded to these items on a 6-point scale ranging from strongly disagree to strongly agree. The scale has been shown to possess good convergent and discriminant validity (Neuberg & Newsom, 1993).

Locus of control. To examine how locus of control is related to the action control mode during sequence production we administered the Internal, Powerful Others, and Chance Locus of Control scales (Levenson, 1981). This 24-item questionnaire differentiates between two types of external orientation: beliefs in the random nature of the world, and beliefs in the predictability of the world coupled with the expectancy that powerful others are in control.

Trajectory SRT. SRT performance was assessed using the trajectory SRT task from (Kachergis et al., 2014), which is a mouse-tracking adaptation of Nissen and Bullemer's SRT task (Nissen & Bullemer, 1987). Similar to the original task, it utilizes 4 different button locations, however, these locations are coded to the corners of a computer screen. As such, the participant has to move the mouse cursor to the targets instead of pressing physical buttons. The

stimuli consisted of 4 red squares (location 1 = upper left, 2 = upper right, 3 = lower left, 4 = lower right), which were displayed continuously throughout the task. Each square was 80 x 80 pixels in size and separated from each other by 440 pixels of white space (Figure 2). Participants were instructed to quickly and accurately move the mouse cursor to whichever square turned green. Shortly after highlighting the green square, the square's color would change back to red, and another square turned green after a 500 ms inter stimulus interval. In the first part of the task participants completed 80 training trials, each containing a sequence of 10 locations. In order to prevent carryover effects between this task and the reinforcement learning task, the present study used a different sequence (3-2-4-2-1-4-3-4-2-1) than the Nissen and Bullemer (1987) study. Every 20 training trials participants were given the opportunity to take a short break of 1 minute. In order to measure the amount of sequence learning, the training phase was followed by a production phase in which the participants were asked to attempt to reproduce any sequence they had previously learned. A correct reproduction of the sequence would not lead to a color change in the squares, while an incorrect reproduction would cause the correct continuation of the sequence to show up in a green color. In order to examine frequency effects the sequence was arranged such that straight movements were more frequent than diagonal movements. At the conclusion of the trajectory SRT task, the participants completed a questionnaire that tested their explicit knowledge. The first question stated "Have you noticed any repeated sequence?". In the scenario that indeed a repeated sequence was noticed, a second question followed: "Can you write down the sequence of locations?". A dichotomous knowledge factor was created post-hoc on the basis of these answers and their correctness. We continually tracked the mouse cursor's x and y pixel coordinates, using which we calculated two measures on which our trajectory analyses are based: (a) predictive movements and (b) correct predictive movements. For (a), we

computed the distance traveled (in pixels) from the previous target before arriving at the next

target during the 500 ms inter stimulus interval. Although these movements can be considered

predictive, they can also simply be incorrect. Considering this, for (b) we computed the distance

(in pixels) between the mouse cursor and the location of the next target. This means that if

participants made a perfect correct predictive mouse movement, the distance would be 0.

      Reinforcement learning. Reinforcement learning performance was assessed using an

adapted version of the trajectory SRT task (see above). The task was adapted to no longer

provide cues regarding the next target location, forcing participants to explore the target options

until the correct target was detected. Participants were presented with target squares in the

corners of the computer screen, which they were instructed to explore using the mouse cursor.

The goal was to maximize the score on the scoreboard, which was located at the top of the screen

and updated with each progression in the task. No indications were provided regarding the

targets' validness as each square was colored blue. Upon reaching a valid target, its color would

shortly change to green and the score would increase by 1 point. Contrarily, reaching an invalid

target would cause the color to change to red, relocating the cursor to the previously (correctly)

occupied target, and penalize the participant by decreasing the score by 1 point. Target validness

was determined by the recurring sequence taken from the Nissen and Bullemer (1987) study (i.e.

4-2-3-1-3-2-4-3-2-1). Unbeknownst to the participants, only one of the target squares would be

valid at any given time and each trial consisted of a sequence of 10 targets labeled 1-4 (location 1

= upper left, 2 = upper right, 3 = lower left, 4 = lower right) that repeated until 80 sequence

iterations were made. No indication where each trial started and ended was provided.

Consequently, in the scenario a participant had a complete understanding of the underlying

sequence before starting and never reached an invalid target, the maximum score of 800 points

would be reached. Contrarily, a participant could theoretically make an infinite number of

mistakes if (s)he had no memory of even the previously attempted target, preventing the

experiment of ever coming to an end. At the conclusion of the reinforcement learning task, the

participants completed a questionnaire that tested their explicit knowledge. The first question

stated "Have you noticed any repeated sequence?". In the scenario that indeed a repeated

sequence was noticed, a second question followed: "Can you write down the sequence of

locations?". A dichotomous knowledge factor was created post-hoc on the basis of these

answers and their correctness.

Design

All participants completed all tasks. To control for order effects, the order in which the

trajectory SRT task and the reinforcement learning task were administered was counterbalanced

over participants. Only the experiment leader was aware of the group to which each participant

was assigned. All participants were given course credit as compensation for their time and

effort.

Procedure

The experiment took place in a laboratory based in the Faculty of Social and Behavioural

 Sciences at Leiden University. After providing written informed consent, participants were

seated in front of a computer monitor after which the experimenter first administered the

Personal Need for Structure questionnaire, followed by the Levenson Multidimensional Locus of

Control questionnaire, the visuospatial working memory task, and Raven's Standard Progressive

Matrices. Following a 5-minute break, participants completed the trajectory SRT task and

reinforcement learning task (order counterbalanced).

Results

Trajectory SRT task

      Data preparation and inspection. In order to prepare the data for further processing, we identified outliers as mouse movement times greater than 1500 ms, which were excluded from further analyses. Trajectory SRT data was split into 10 blocks of 8 sequence in the block factor, allowing us to examine developments in sequence learning by comparing task accuracy and mouse movement times across blocks. Furthermore, sensitivity to sequence frequency information was examined by distinguishing between straight (frequent) and diagonal (infrequent) mouse movements in the movement type factor.

      Movement times and accuracy. Developments in mouse movement times were examined comparing participant's median mouse movement times across blocks. Mouse movement time was defined as the time between trial onset (stimulus square turning green) and trial end (mouse cursor touching the stimulus). Median mouse movement time was 464 ms (*SD* = 222.53). Analyses revealed participants became faster over time, replicating the Nissen and Bullemer (1987) movement time results, with a mean movement time of 592 ms in the first block to 496 ms in the final block, $F(9, 351) = 15.51$, $p < .001$ (Figure 3). However, this speed-up went together with a significantly decrease in accuracy over time, $F(3.78, 147.42) = 4.43$, $p < .01$, suggesting a speed-accuracy trade-off among participants (Figure 4). Mauchly's test (Mauchly, 1940) indicated that the assumption of sphericity had been violated for the block factor ($W < .001$, $p < .01$), therefore, the degrees of freedom and *p*-value have been corrected using Huynh-Feldt estimate of sphericity ($\varepsilon = .42$; Huynh & Feldt, 1976).

      Sequence knowledge. Depending on whether participants had explicit knowledge of the underlying sequence pattern, participants were classified as either implicit or explicit learners in

the knowledge factor. The percentage of implicit learners was 67.5% (27 out of 40 participants),

and the percentage of explicit learners was 32.5% (13 out of 40 participants). Participants with

explicit sequence knowledge showed to have a significantly larger visuospatial working memory

capacity ($M = 2.87$) as compared to participants without explicit knowledge ($M = 2.25$), $t(28.08)$

$= 2.95$, $p < .01$. No significant differences were found between implicit and explicit learners on

estimated IQ ($t(24.99) = -.55$, $p = .59$), the Levenson Multidimensional Locus of Control scales,

$t(35.40) = .81$, $p = .42$, and Personal Need for Structure scales, $t(22.28) = .29$, $p = .77$.

Predictive movements. As described in the Materials and measures section above,

developments in predictive movements was examined by computing the distance traveled from

the previous target before arriving at the next target during the inter stimulus interval. Data were

analyzed using an analysis of variance (ANOVA) with a within-subjects factor of block. The

results show that over time participants traveled longer distances with their mouse cursors during

the inter stimulus interval, with an average of 176.91 pixels traveled in block 1, to an average of

311.13 pixels in block 10, $F(3.87, 150.93) = 4.85$, $p < .01$ (Figure 5). An ANOVA on predictive

movements with knowledge as the between-subject and block as the within-subject factors,

showed a significant main effect of block, $F(4.05, 153.9) = 6.53$, $p < .001$, but not for

knowledge, $F(1, 38) = .15$, $p = .69$. Mauchly's test indicated that the assumption of sphericity

had been violated for the block factor in both analyses of variance ($Ws < .001$, $ps < .01$),

therefore, the degrees of freedom and p-values have been corrected using Huynh-Feldt estimate

of sphericity ($\varepsilon = .43$ and $\varepsilon = .45$, respectively). Visual inspection did not suggest an association

between predictive movements and either visuospatial working memory capacity, estimated IQ,

the Levenson Multidimensional Locus of Control scales, or the Personal Need for Structure

scales.

Correct predictive movements. Mouse movements towards the next target in the sequence, prior to the target's appearance, are regarded as *correct* predictive movements and to be also indicative of the current action control mode employed. Similar to the previous analysis, an ANOVA with a within-subjects factor of block was performed. As sequence learning progressed, participants shortened the distance between the mouse cursor and the next target, suggesting a shift from stimulus-based to plan-based control, with an average of 607.5 pixels distance in the first block, to an average of 469.64 pixels distance in the final block, $F(3.33, 129.87) = 16.52$, $p < .001$ (Figure 6). An ANOVA on correct predictive movements with knowledge as between-subject and block as within-subject factors, showed significant main effects of block, $F(4.68, 177.84) = 32.26$, $p < .001$, and knowledge, $F(.52, 19.76) = 12.95$, $p < .001$. The significant block * knowledge interaction, $F(4.68, 177.84) = 14.00$, $p < .001$, revealed that explicit learners, in contrast to implicit learners, strongly progressed in reducing the distance between the mouse cursor and the next target during the inter stimulus intervals (Figure 7). Mauchly's test indicated that the assumption of sphericity had been violated for the block factor in both analyses of variance (Ws $< .001$, ps $< .01$), therefore, the degrees of freedom and p-values have been corrected using Huynh-Feldt estimate of sphericity ($\varepsilon = .37$ and $\varepsilon = .52$, respectively). Similar to the previous section, visual inspection did not suggest an association between predictive movements and either visuospatial working memory capacity, estimated IQ, the Levenson Multidimensional Locus of Control scales, or the Personal Need for Structure scales.

Frequency effects. No differences were found between participants in the first block and the final block concerning sensitivity to frequency information, $t(78) = .26$, $p = .80$, suggesting no general trend of participants shifting from stimulus-based to plan-based control (see Figure 8; cf. Hoffmann & Koch, 1997). In order to test our hypothesis whether action plan formation over

time is dependent on the development of explicit sequence knowledge, an ANOVA on mouse

movement times with knowledge as between-subject, and block and movement type as within-

subject factors was performed. It was expected that both types of learners start with a similar

level of sensitivity to frequency information, but that over time explicit learners develop an

action plan, and thus declining frequency effect. Our results suggest that this is not the case

(Figure 9), evidenced by significant main factors and two-way interactions, but not the necessary

three-way interaction between block, movement type, and knowledge, $F(9, 342) = .79, p = .63$

(see Table 1). Post-hoc analyses revealed that frequent (straight) mouse movements were

executed faster ($M = 417.27$ ms, $SD = 105.43$ ms) than infrequent (diagonal) mouse movements

($M = 495.93$ ms, $SD = 115.92$ ms). Participants with explicit sequence knowledge had shorter

mouse movement times ($M = 398.2$ ms, $SD = 120.85$ ms) as compared to participants without

explicit sequence knowledge ($M = 484.71$ ms, $SD = 90.42$ ms).

Reinforcement learning task

      Data inspection and preparation. Similar to the trajectory SRT task, depending on

whether participants had explicit knowledge of the underlying sequence pattern, participants

were classified as either explicit learners or implicit learners. The percentage of implicit learners

was 42.5% (17 out of 40 participants), and for explicit learners the percentage was 57.5% (23 out

of 40 participants). Inspection of the distribution of reinforcement learning task scores revealed

that scores were non-normally distributed, with a group of participants scoring around 700 points

and almost all remaining participants scoring relatively low (Figure 10). Indeed, the Shapiro-

Wilk test (Shapiro & Wilk, 1965) revealed that the hypothesis of normality could be rejected, W

$= .93, p < .02.$ Even though Hartigans' dip test (Hartigan & Hartigan, 1985) was unable to

provide additional evidence for a bimodal distribution of task scores ($D = .038$, $p = .952$), due to the convincing indications of the presence of two clusters of scores a mid-range split on 457 points was performed, allowing us to distinguish between low and high performers.

Determinants of score and sequence knowledge. Neither mouse movement times nor sequence knowledge in the trajectory SRT task were found to be predictive for reinforcement learning scoring, $t(31.75) = 1.04$, $p = .30$, and $t(29.15) = -1.41$, $p = .17$, respectively. However, trajectory SRT task sequence knowledge was found to significantly predict sequence knowledge in the reinforcement learning task, $X^2(1) = 4.5$, $p < .05$, suggesting that sequence learning for both tasks is dependent on similar constructs, or that the discovery of a repeating sequence in the second-last task (trajectory SRT or reinforcement learning task) might have made participants cautious for the detection of a hidden sequence in the final task. Knowledge of the underlying sequence in the reinforcement learning task was found to be predictive of reinforcement learning score, with explicit learners reaching a much higher score ($M = 634.05$) than implicit learners ($M = 374.65$), $t(24.67) = -4.61$, $p < .001$.

When distinguishing between high and low performing participants, high performers had a significantly higher mean estimated IQ of 104.36, compared to low performers with an estimated mean of 91.91, $t(38) = -2.70$, $p < .05$ (Figure 11). Similarly, high performers on the reinforcement learning task had a higher mean visuospatial working memory capacity of 2.63, compared to low performers with a mean capacity of 2.12, $t(38) = -2.29$, $p < .05$ (Figure 12). No significant differences were found between high and low performers on the Levenson Multidimensional Locus of Control scales, $t(38) = -.50$, $p = .62$, and Personal Need for Structure questionnaire, $t(38) = .10$, $p = .92$, suggesting that sequence learning is not dependent on personality characteristics, but rather on cognitive capabilities.

Discussion

Using novel trajectory adaptations of the SRT task, the present study examined for the first time whether cognitive capabilities and personality characteristics are determinants of one's action control mode (stimulus-based and plan-based). It was hypothesized that differences in sequence learning could be attributed to differences in visuospatial working memory capacity, IQ, locus of control, personal need for structure, and sequence knowledge acquisition. Further, the present study aimed at reproducing the bimodal score distribution found in the SRT reinforcement learning paradigm in Kachergis et al. (2016) and examine for the first time whether the distinctive peaks are a product of different control modes.

By implementing a trajectory adaptation of the SRT paradigm (Kachergis et al., 2014) we were able to examine the sequence learning process with higher granularity than the original SRT paradigm. Whereas Nissen and Bullemer (1987)'s original SRT task is limited to measuring keypresses, the trajectory adaptation offers the advantage of being able to record mouse movements toward the next predicted stimulus. Because plan-based control is characterized by making predictive movements in the absence of cues (Nattkemper & Prinz, 1997) detection of correct predictive movements was considered to be indicative of the current action control mode employed. Our results show that signs of sequence learning were evident by a steady increase in predictive movements. Further, as plan-based control is thought to rely on an action plan (Hommel, 2003; Luria, 1961; Miller et al., 1960) which can be expressed explicitly (Tubau et al., 2007), we used the ability to verbally report the sequence learned as an additional indicator of plan-based control. As not all participants were able of correctly report the sequence learned, we were able to distinguish between implicit and explicit learners and examine possible differences in their cognitive capabilities and personality characteristics. This allowed us to shed light on

whether these individual differences are indicative for the presence or absence of a shift in executive control mode. Indeed, we found explicit sequence knowledge to be predicted by visuospatial working memory in the trajectory SRT task. In earlier work, Bo et al. (2011) found visuospatial working memory to be predictive of SRT performance. In the present study we found this also to be case, as visuospatial working memory was predictive of sequence knowledge in both the trajectory and reinforcement learning adaptations of the SRT task. Finally, we used a sequence in which straight movements were more frequent than diagonal movements to examine sensitivity to frequency information, which was found by Tubau et al. (2007) to be diminished under plan-based control. We argued that action plan formation could be dependent on the development of explicit sequence knowledge. Specifically, we expected that both implicit and explicit learners start with similar levels of sensitivity to frequency information, but that over time explicit learners would demonstrate declining frequency effect. In spite of accounting for the different types of learning among participants, no differences were found between implicit and explicit learners in sensitivity to frequency information. In the reinforcement learning task, we found a non-normal distribution of scores and a cluster of high performing participants similar to what has been found in Kachergis et al. (2016). When distinguishing between low and high performing participants, high performers showed to have higher visuospatial working memory capacity and estimated IQ (Figures 11 and 12), but did not differ from low performers in locus of control and personal need for structure. This implies that exploration-based sequence learning is determined by one's cognitive capabilities, rather than personality characteristics.

It is important to note that our current experimental design was not without shortcomings. The present study implemented a sequence with straight and diagonal movements, of which the

former were more frequent than the latter. Our failed attempt to reproduce the frequency effects observed in Tubau et al. (2007) could be explained by the increased number of dimensions in our implemented trajectory SRT paradigm. Whereas the SRT paradigm in Tubau et al. (2007) allows for only horizontal repetitions and switches, its trajectory adaptation allows for many more movements to be made (horizontal repetitions and switches, vertical repetitions and switches, and diagonal repetitions and switches). Consequently, numerous frequency effects could have been active, lowering the usability of using sensitivity to frequency information as an indicator of plan-based control in this paradigm. Moreover, an additional shortcoming of this paradigm is ingrained by the placement of the visual stimuli in the corners of a computer screen. While this allows for distinguishing between straight and diagonal movements, it results in unequal distances between target pairs as diagonal movements have longer distances than straight movements, possibly making diagonal movements a less appealing alternative.

Even though mouse-tracking SRT paradigms allow for revealing novel information about ongoing motor execution and learning processes, they do not offer mechanistic explanations for sequence learning. More sophisticated approaches are needed to infer properties of sequence acquisition mechanisms. One such approach is computational modeling, which uses computers to simulate and study the behavior of complex systems using mathematics. In computational modeling theories are formalized in such a way that they can be implemented as computer programs. Consequently, it offers the possibility of gaining insights in our theoretical predictions and inferring properties of sequence learning mechanisms by examining what model parameters are responsible for particular outcomes. Indeed, computational modeling has become a well-established approach in many disciplines, including cognitive science (Pylyshyn, 1984) and artificial

intelligence (Partridge & Wilks, 1990). With regards to the current topic of investigation, computational models could help us better understand the mechanisms underlying sequence acquisition in a reinforcement learning paradigm. For instance, future research could fit a reinforcement learning model to human data and examine how visuospatial working memory capacity and IQ are related to the model's parameters, gaining valuable insights in the mechanisms underlying sequential action acquisition in an exploration-based paradigm.

To conclude, we believe that the combination of both cued sequential action and exploration-based paradigms gave us valuable insights in the mechanisms underlying reinforcement learning. Our results suggest that the link between determinants of action control modes and reinforcement learning performance is not clear yet, and that future research is needed to gain a better understanding of the conditions under which frequency effects can indeed be used as indicators of action control mode.

References

Bo, J., Jennett, S., & Seidler, R. D. (2011). Working memory capacity correlates with implicit

    serial reaction time task performance. *Experimental Brain Research*, *214*(1), 73–81. doi:

    10.1007/s00221-011-2807-8

Cleeremans, A., & McClelland, J. L. (1991). Learning the structure of event sequences. *Journal*

    of Experimental Psychology: General, *120*(3), 235–253. doi: 10.1037/0096-3445.120.3.235

Cohen, A., Ivry, R. I., & Keele, S. W. (1990). Attention and structure in sequence learning.

    Journal of Experimental Psychology: Learning, Memory, and Cognition, *16*(1), 17–30. doi:

    10.1037/0278-7393.16.1.17

Destrebecqz, A., & Cleeremans, A. (2001). Can sequence learning be implicit? New evidence

    with the process dissociation procedure. *Psychonomic Bulletin & Review*, *8*(2), 343–350.

    doi: 10.3758/BF03196171

Dienes, Z., Broadbent, D. E., & Berry, D. C. (1991). Implicit and explicit knowledge bases in

    artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and*

    Cognition, *17*(5), 875–887. doi: 10.1037/0278-7393.17.5.875

Hartigan, J. A., & Hartigan, P. M. (1985). The dip test of unimodality. *The Annals of Statistics*,

    13(1), 70–84. doi: 10.1214/aos/1176346577

Hazeltine, E. (2002). The representational nature of sequence learning: Evidence for goal-based

    codes. In *Common mechanisms in perception and action: Attention and performance XIX*

    (pp. 673–690). Oxford, UK: Oxford University Press.

Herwig, A., Prinz, W., & Waszak, F. (2007). Two modes of sensorimotor integration in

    intention-based and stimulus-based actions. *Quarterly Journal of Experimental Psychology*.

    60(11), 1540–1554.. doi: 10.1080/17470210601119134

Hoffmann, J., & Koch, I. (1997). Stimulus-response compatibility and sequential learning in the

    serial reaction time task. *Psychological Research*, *60*(1–2), 87–97. doi:

    10.1007/BF00419682

Hommel, B. (2003). Planning and representing intentional action. *The Scientific World Journal*,

    3, 593–608. doi: 10.1100/tsw.2003.46

Huynh, H., & Feldt, L. S. (1976). Estimation of the Box correction for degrees of freedom from

    sample data in randomized block and split-plot designs. *Journal of Educational Statistics*,

    1(1), 69–82. doi: 10.2307/1164736

James, W. (1890). *The principles of psychology (Vol. 2)*. New York, US: Holt. doi:

    10.1037/10538-000

Jiménez, L., & Méndez, C. (1999). Which attention is needed for implicit sequence learning?

    Journal of Experimental Psychology: Learning, Memory, and Cognition, *25*(1), 236–259.

    doi: 10.1037/0278-7393.25.1.236

Kachergis, G., Berends, F., de Kleijn, R., & Hommel, B. (2016). Human reinforcement learning

    of sequential action. *Proceedings of the 38th Annual Conference of the Cognitive Science

    Society*, (pp. 193-198). Retrieved from https://mindmodeling.org/cogsci2016/papers/0046/

Kachergis, G., Berends, F., Kleijn, R. de, & Hommel, B. (2014). Trajectory effects in a novel

    serial reaction time task. *Proceedings of the 36th Annual Conference of the Cognitive

    Science Society*, (pp. 713-718). Retrieved from

    https://mindmodeling.org/cogsci2014/papers/131/

Knowlton, B. J., & Squire, L. R. (1994). The information acquired during artificial grammar

    learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(1),

    79–91. doi: 10.1037/0278-7393.20.1.79

Knowlton, B. J., & Squire, L. R. (1996). Artificial grammar learning depends on implicit

   acquisition of both abstract and exemplar-specific information. *Journal of Experimental

   Psychology: Learning, Memory, and Cognition*, *22*(1), 169–181. doi: 10.1037/0278-

   7393.22.1.169

Levenson, H. (1981). Differentiating among internality, powerful others, and chance. *Research

   with the Locus of Control Construct*, 15–63. doi: 10.1016/B978-0-12-443201-7.50006-3

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and

   conjunctions. *Nature*, *390*(6657), 279–281. doi: 10.1038/36846

Luria, A. R. (1961). *The role of speech in the regulation of normal and abnormal behavior*. New

   York, US: Cambridge University Press. doi: 10.1016/c2013-0-05275-9

Martini, M., Furtner, M. R., & Sachse, P. (2013). Working memory and its relation to

   deterministic sequence learning. *PLoS ONE*, *8*(2), e56166. doi:

   10.1371/journal.pone.0056166

Mauchly, J. W. (1940). Significance test for sphericity of a normal n-variate distribution. *The

   Annals of Mathematical Statistics*, *11*(2), 204–209. doi: 10.1214/aoms/1177731915

Meulemans, T., Van der Linden, M., & Perruchet, P. (1998). Implicit sequence learning in

   children. *Journal of Experimental Child Psychology*, *69*(3), 199–221. doi:

   10.1006/JECP.1998.2442

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior.* New

   York, US: Holt. doi: 10.1525/aa.1960.62.6.02a00190

Münsterberg, H. (1892). *Beiträge zur Experimentellen Psychologie (Heft IV) [Contributions to

   experimental psychology (Issue IV)]*. Freiberg, Germany: Mohr.

Nattkemper, D., & Prinz, W. (1997). Stimulus and response anticipation in a serial reaction task.

Psychological Research, *60*(1–2), 98–112. doi: 10.1007/BF00419683

Neuberg, S. L., & Newsom, J. T. (1993). Personal need for structure: Individual differences in the desire for simpler structure. *Journal of Personality and Social Psychology*, *65*(1), 113–131. doi: 10.1037/0022-3514.65.1.113

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, *19*(1), 1–32. doi: 10.1016/0010-0285(87)90002-8

Partridge, D., & Wilks, Y. (1990). *The foundations of artificial intelligence: A sourcebook*. New York, US: Cambridge University Press. doi: 10.1017/cbo9780511663116

Pylyshyn, Z. W. (1984). *Computation and cognition: Toward a foundation for cognitive science*. Cambridge, US: MIT Press. doi: 10.1002/bs.3830310408

Rauch, S. L., Savage, C. R., Brown, H. D., Curran, T., Alpert, N. M., Kendrick, A., … Kosslyn, S. M. (1995). A PET investigation of implicit and explicit sequence learning. *Human Brain Mapping*, *3*(4), 271–286. doi: 10.1002/hbm.460030403

Raven, J. C., Court, J. H., & Raven, J. (1998). *Manual for Raven's progressive matrices and vocabulary scales. Section 4: The advanced progressive matrices.* San Antonio, US: Harcourt Assessment.

Raven, J., Raven, J. C., & Court, J. H. (2000). *Manual for Raven's standard progressive matrices and vocabulary scales: Section 3: The standard progressive matrices.* Oxford, UK: Oxford Psychologists Press.

Schendan, H. E., Searl, M. M., Melrose, R. J., & Stern, C. E. (2003). An fMRI Study of the role of the medial temporal lobe in implicit and explicit sequence learning. *Neuron*, *37*(6), 1013–1025. doi: 10.1016/S0896-6273(03)00123-5

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete

samples). *Biometrika*, *52*(3–4), 591–611. doi: 10.1093/biomet/52.3-4.591

Thompson, M. M., Naccarato, M. E., & Parker, K. (1989). Assessing cognitive need: The

development of the personal need for structure and personal fear of invalidity scales.

Proceedings of the Annual Meeting of the Canadian Psychological Association. Halifax,

Canada.

Tubau, E., Hommel, B., & López-Moliner, J. (2007). Modes of executive control in sequence

learning: From stimulus-based to plan-based control. *Journal of Experimental Psychology:

General*, *136*(1), 43–63. doi: 10.1037/0096-3445.136.1.43

Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in

visual working memory capacity. *Nature*, *428*(6984), 748–751. doi: 10.1038/nature02447

Vygotsky, L. (1986). *Thought and language*. Cambridge, US: MIT Press. doi:

10.1017/S0272263100008172

Zelazo, P. D. (1999). Language, levels of consciousness, and the development of intentional

action. In P. D. Zelazo, W. Astington, & D. R. Olson (Eds.), *Developing theories of

intention: Social understanding and self-control* (pp. 95–117). Mahwah, US: Erlbaum.

Table 1.

Table 1.

Analysis of variance of block, knowledge, and movement type on mouse movement times.

Factor	df	F	η^2_G	p
Block[a]	5.31, 201.78	17.52	.129	< .001
Knowledge	1, 38	6.44	.089	< .05
Movement type	1, 38	104.12	.084	< .001
Block * Knowledge[a]	5.31, 201.78	8.37	.066	< .001
Block * Movement type	9, 342	2.7	.005	< .01
Knowledge * Movement type	1, 38	4.43	.004	< .05
Block * Knowledge * Movement type	9, 342	.79	.001	.63

[a]Mauchly's test (Mauchly, 1940) indicated that the assumption of sphericity had been violated (Ws < .01, ps < .001), therefore, the degrees of freedom and p-value have been corrected using Huynh-Feldt estimate of sphericity ($\varepsilon = .59$).

Figure 1.
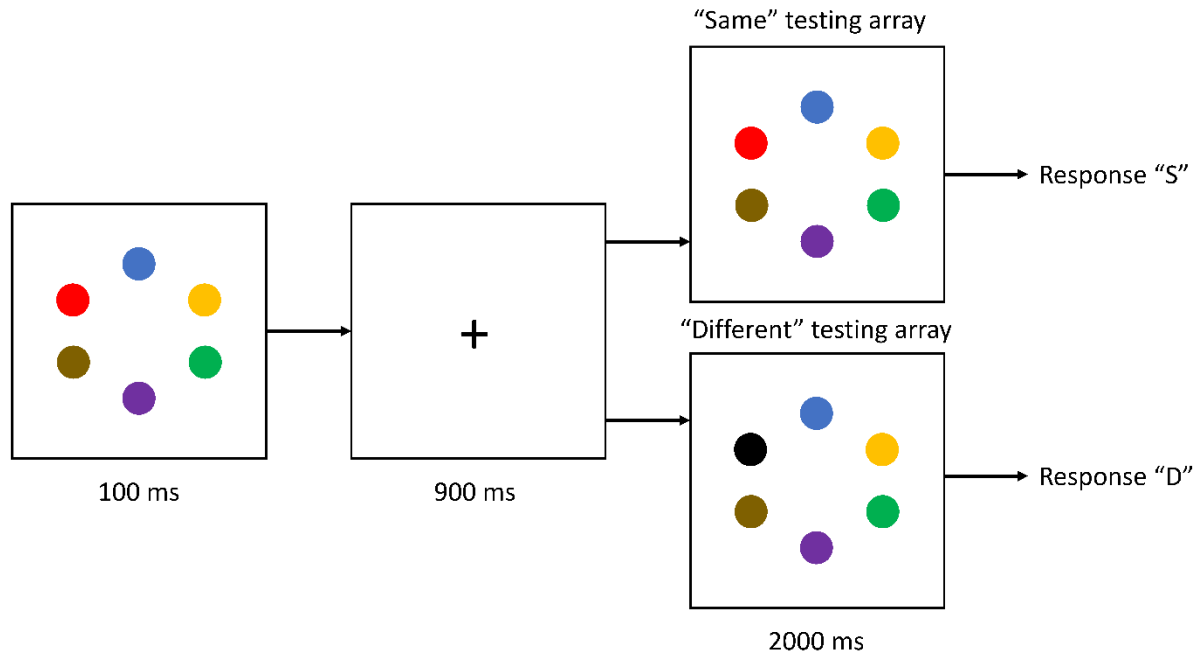
Figure 1. Illustration of the visuospatial working memory task from Bo et al. (2011). The sample array consisted of 2-8 colored circles, presented in varying colors (red, orange, yellow, green, blue, violet, pink, white, black, and brown) on a white background. In each trial, the test array was either identical to the sample array or different with only one of the colors changed. Participants were instructed to indicate whether the test array was the same (response 'S') or different (response 'D') from the sample array by keypress.

Figure 2.
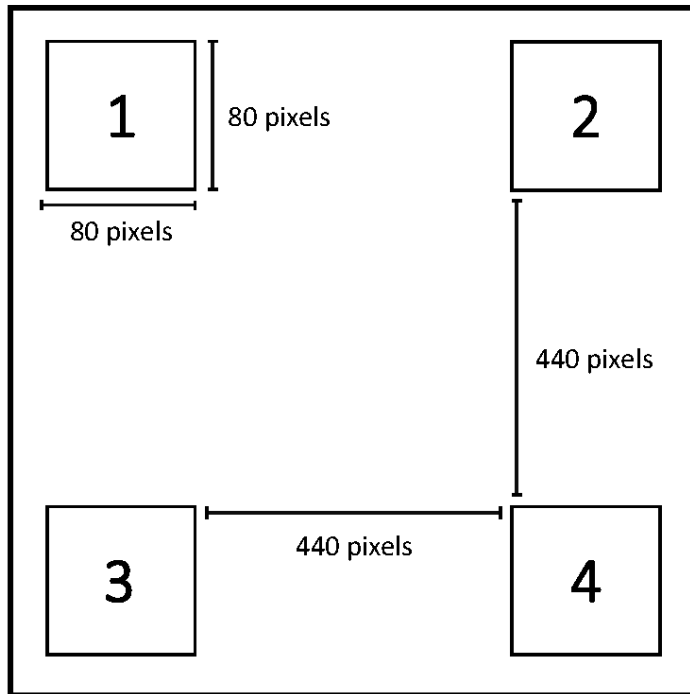
Figure 2. Layout of the trajectory SRT task (Kachergis et al., 2014). Instead of measuring

discrete button-presses as in the original SRT task (Nissen & Bullemer, 1987), its trajectory

adaptation requires participants to respond to stimulus changes by a corresponding move of the

mouse cursor.

Figure 3.
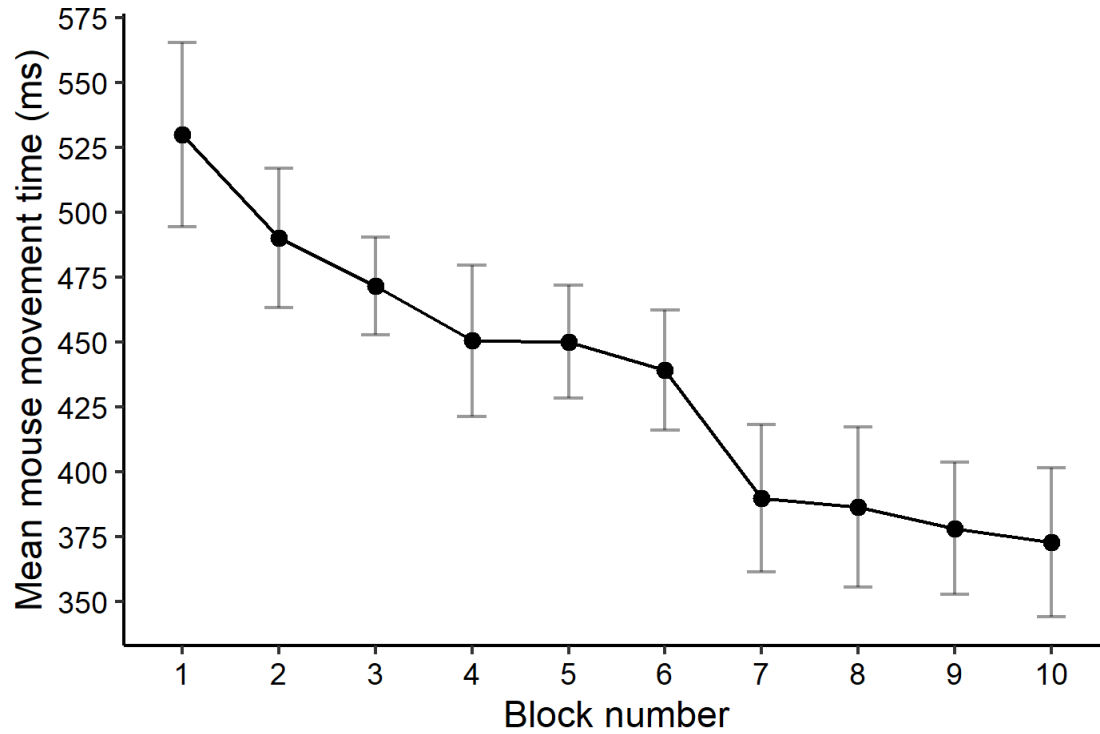
Figure 3. Progression of participants' mean mouse movement time across blocks for the trajectory SRT task. The decrease in movement times suggests learning of the underlying sequence. Error bars indicate 95% confidence interval (CI).

Figure 4.
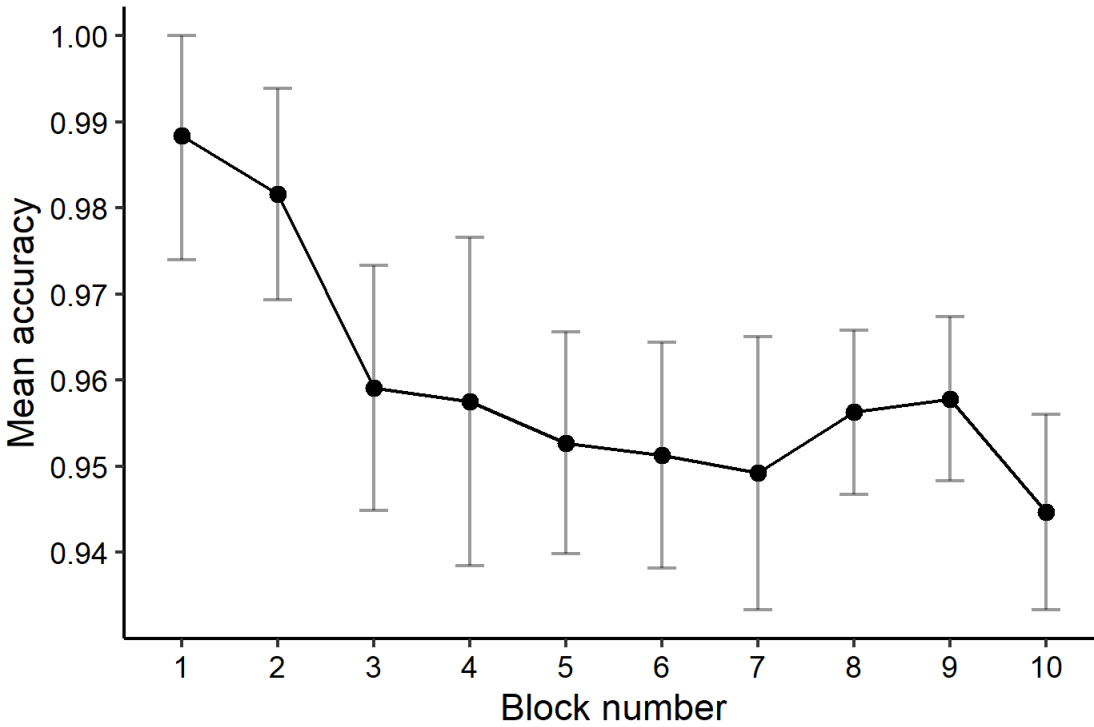
Figure 4. Progression of participants' mean accuracy across blocks in the trajectory SRT task.

When considering the speed increase depicted in Figure 3, the strong decrease in accuracy in the

first three blocks suggests a speed-accuracy trade-off among participants. Error bars indicate

95% CI.

Figure 5.
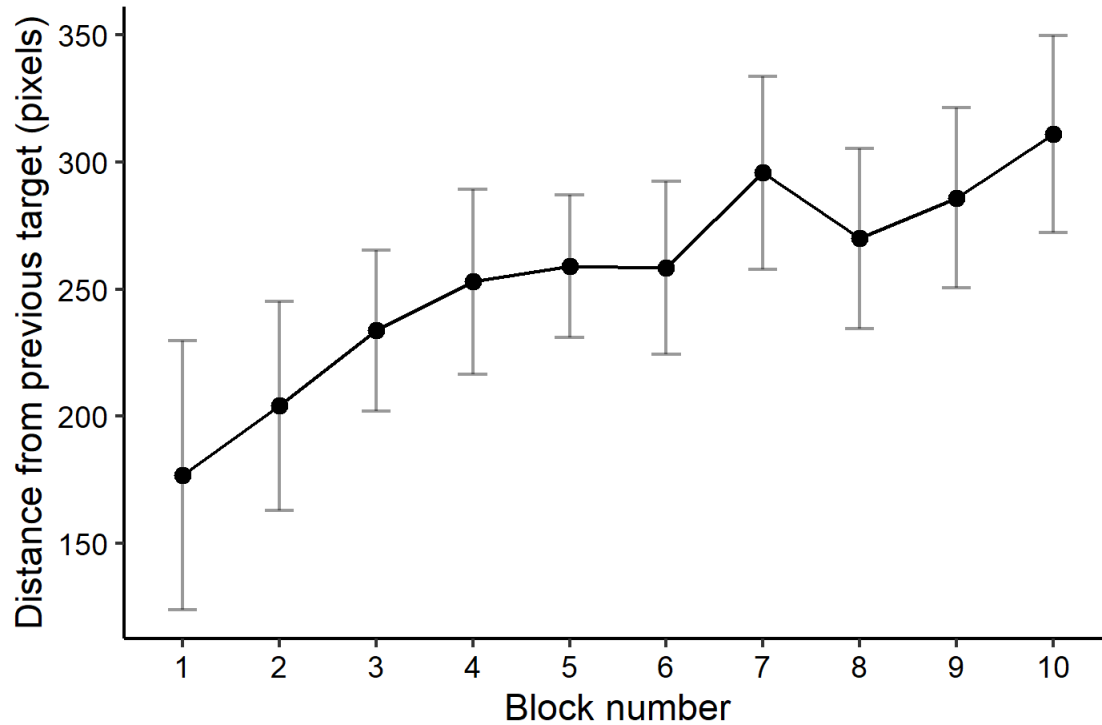
Figure 5. Mean distance traveled (in pixels) from the previous target before arriving at the next

target during the 500-ms inter stimulus interval, per block. Over time participants traveled

longer distances during the inter stimulus interval, suggesting learning of the sequence. Error

bars indicate 95% CI.

Figure 6.
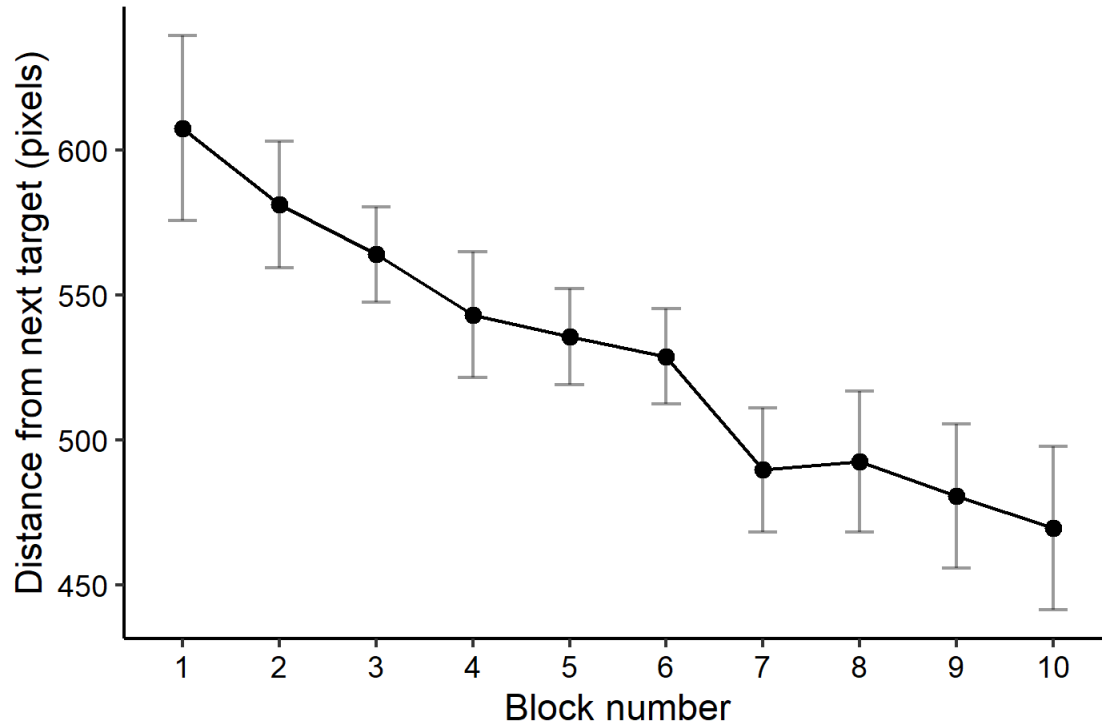
Figure 6. Mean distances between the mouse cursor and the location of the next target (in pixels) during the 500-ms inter stimulus interval, per block. Over time, the distances between the cursor and the next stimulus decreased, suggesting that participants shifted from stimulus-based to plan-based control over time. Error bars indicate 95% CI.

Figure 7.
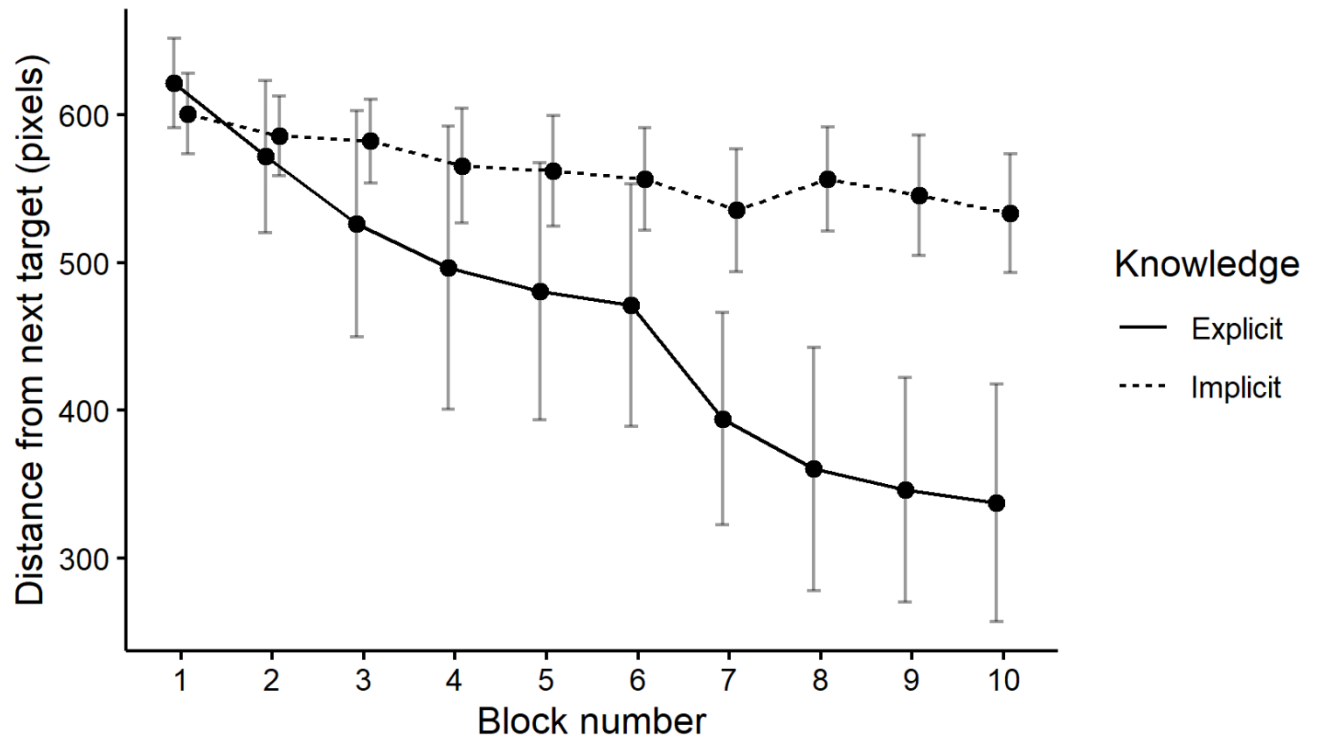
Figure 7. Mean distances between the mouse cursor and the location of the next target during

the inter stimulus interval, split by implicit and explicit sequence knowledge. Over time explicit

learners exhibited significantly more correct predictive responses than implicit learners. Error

bars indicate 95% CI.

Figure 8.
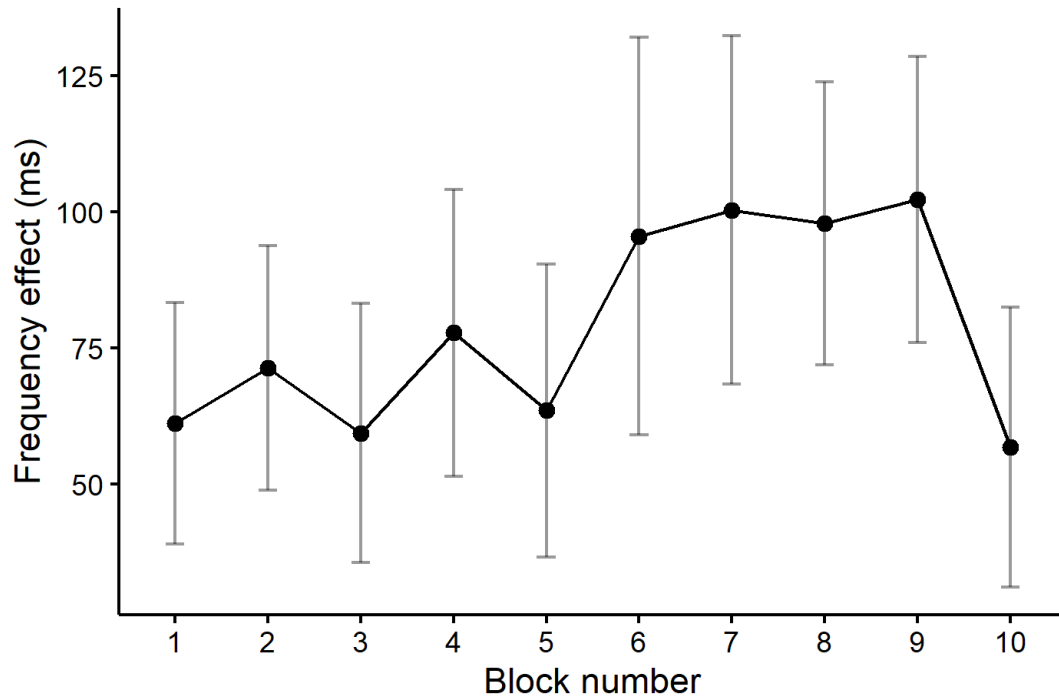
Figure 8. Development of mean frequency effect in the trajectory SRT task, split by block.

Participants became increasingly sensitive to frequency information, with a steep decrease in

sensitivity in the final block. Error bars indicate 95% CI.

Figure 9.
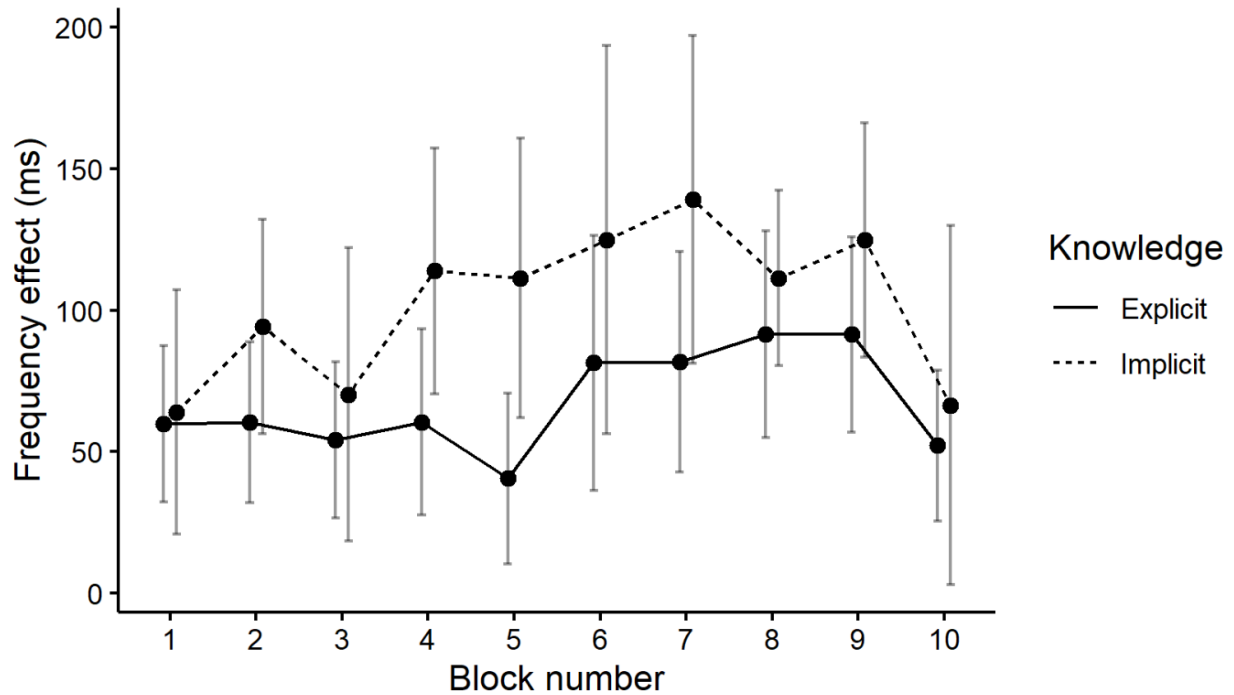
Figure 9. Development of mean frequency effect (i.e. difference in mouse movement times between straight [frequent] and diagonal [infrequent] movements) in the trajectory SRT task across blocks, split by implicit and explicit sequence knowledge. Both implicit and explicit learners started with similar sensitivity to frequency information, which after diverging developments, converged back to each other in the final block. Error bars indicate 95% CI.
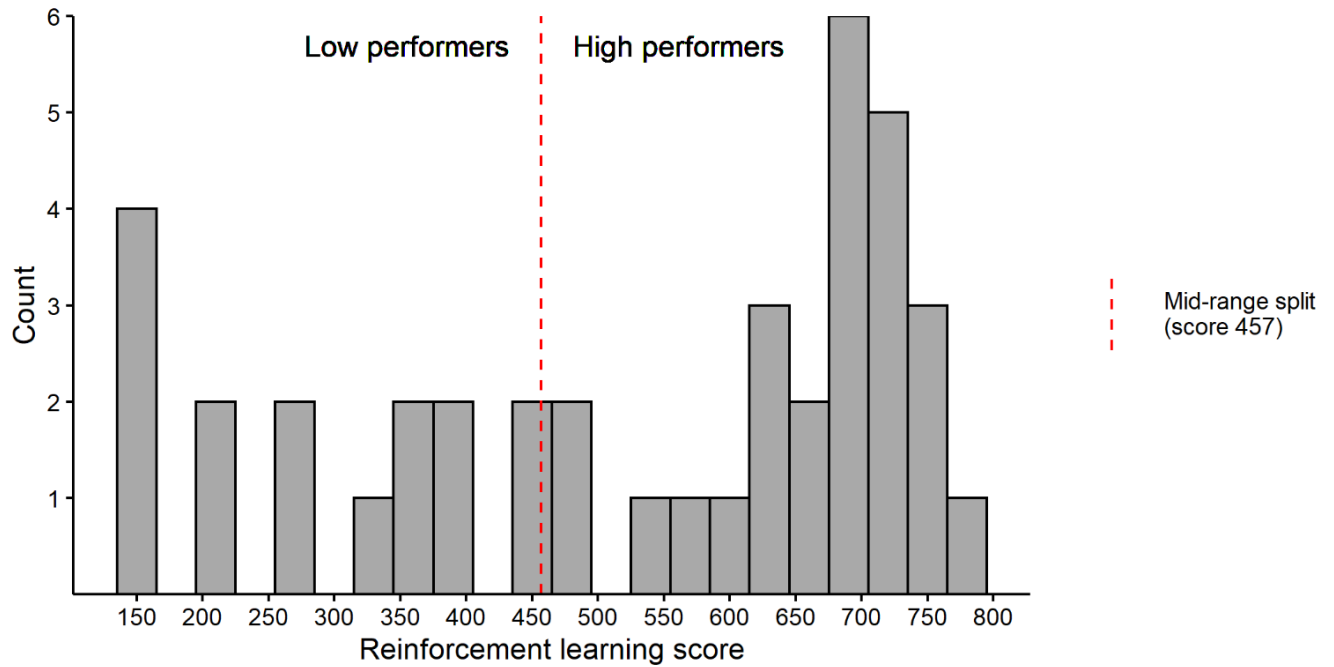
Figure 10.

Figure 10. Distribution of reinforcement learning task scores for all participants. Distribution of scores was non-normal, with a large group of participants scoring 700 points, and a group scoring quite low. In order to examine differences between high and low scoring participants, we distinguished between high and low scoring participants using the mid-range split score of 457 points. The left side of the red dashed line is considered to comprise of low performing participants, while the right is considered to comprise of high performing participants.
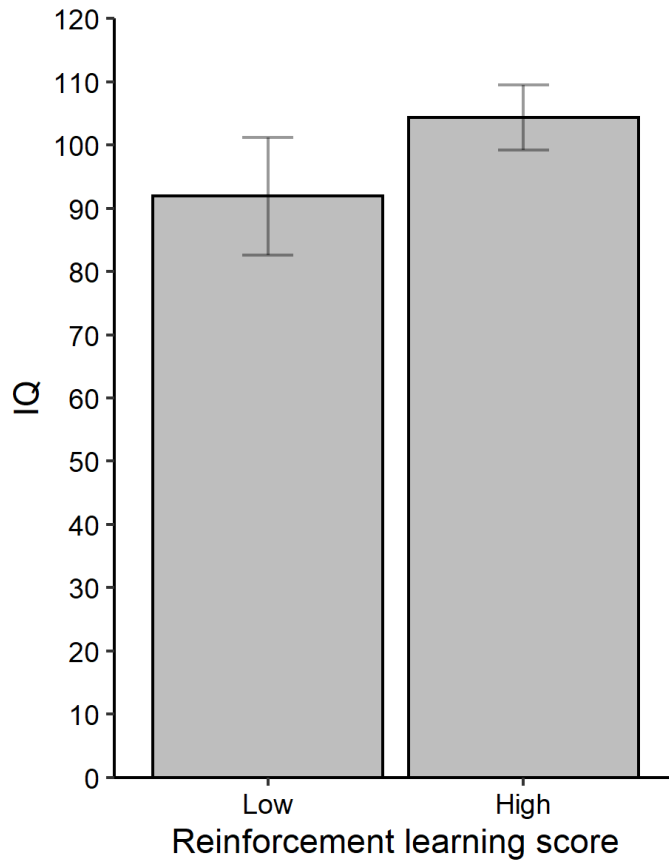
Figure 11.

Figure 11. Mean differences in fluid intelligence, as assessed by Raven's Standard Progressive

Matrices (Raven et al., 1998), between low and high performing participants on the reinforce-
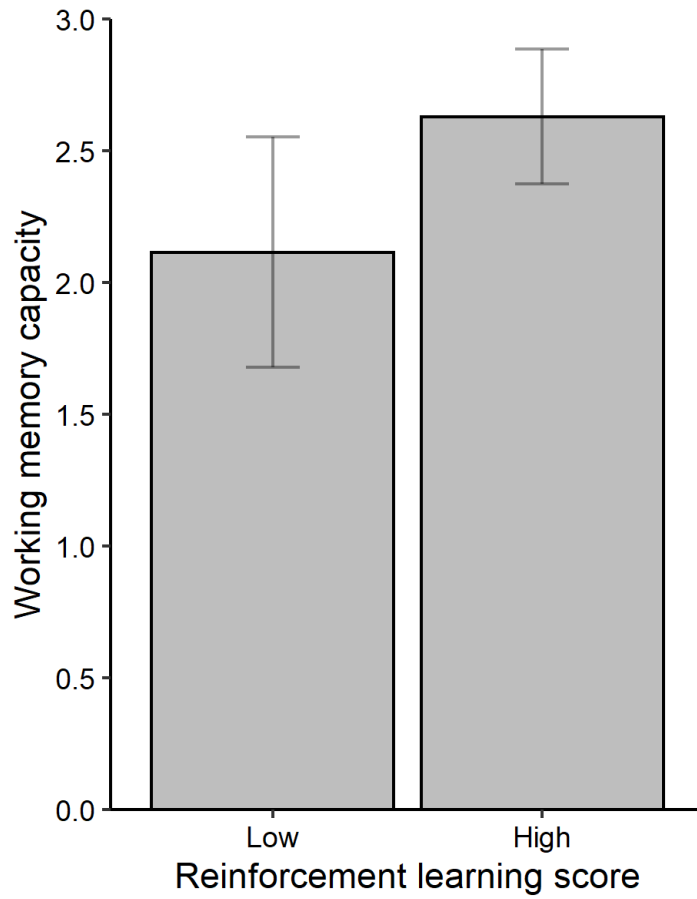
ment learning task. Error bars indicate 95% CI.

Figure 12.

Figure 12. Mean differences in working memory capacity, as assessed by Bo et al. (2011)'s

visuospatial working memory task, between low and high performing participants in the rein-

forcement learning task. Error bars indicate 95% CI.