Acta Psychologica 170 (2016) 168–176

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy

Sequential effects and sequence learning in a three-choice serial reaction time task

Jessica C. Lee [a,*], Tom Beesley [b], Evan J. Livesey [a]

[a] University of Sydney, Sydney, NSW 2006, Australia
[b] University of New South Wales, Sydney, NSW 2052, Australia

ARTICLE INFO

Article history:
Received 5 November 2015
Received in revised form 18 July 2016
Accepted 4 August 2016
Available online xxxx

Keywords:
Sequential effect
Sequence learning
Serial reaction time
Serial recurrent network
Contingency

ABSTRACT

The recent history of events can influence responding despite there being no contingent relationship between those events. These 'sequential effects' are ubiquitous in cognitive psychology, yet their study has been dominated by two-choice reaction time tasks in which sequences necessarily comprise simple response repetitions and alternations. The current study explored sequential effects in a three-choice reaction time task where the target was constrained to either move clockwise or anticlockwise on each trial, allowing for assessment of sequential effects involving the direction of target transitions rather than target location. Across two experiments, a reliable pattern of sequential effects was found in the absence of contingencies, whereby the most notable feature was that participants were fastest to respond to subsequences where the target moved in a consistent direction on consecutive trials, compared to when the target direction alternated. In Experiment 2, the direction of motion was biased to move in one direction 75% of the time and in a subsequent transfer phase, participants showed evidence of learning this probabilistic sequence but still exhibited the same pattern of sequential effects on trials where the target moved in the more prevalent or less prevalent direction. Simulations with a connectionist model of sequence learning (the Augmented Serial Recurrent Network, Cleeremans & McClelland, 1991) produced an adequate replication of the sequential effects in both experiments in addition to an effect of sequence learning in Experiment 2. We propose that sequential effects may represent learning about transient contingencies and may be described using the same associative learning mechanisms intended for sequence learning.

The recent history of events produces noticeable effects on both controlled decision-making (e.g. the gambler's fallacy, Burns & Corpus, 2004; Jarvik, 1951; categorization judgements, Jones, Love, & Maddox, 2006; Jones & Sieck, 2003), as well as more automatic responses (e.g. conditioned responding, Perruchet, 1985; pain sensation, Link, Kos, Wager, & Mozer, 2011). These transient differences in performance as a function of trial history are known as sequential effects, and have been studied most extensively in choice reaction time (RT) procedures, such as the serial reaction time (SRT) task (Nissen & Bullemer, 1987). In this task, participants usually observe a target appearing in one of several locations on the screen, and have to respond with a corresponding keypress. When the task is entirely unstructured, such that there is no consistent sequence to the target's movement between positions, participants are nevertheless faster to respond on certain trials. These "sequential effects" suggest that in the absence of any predictive information, responding is still influenced by recent prior events.

In SRT tasks, sequential effects are normally differentiated from sequence learning. To examine the latter, contingencies are embedded between target locations, such that reductions in RTs for predictable (sequenced) trials reliably occur over the course of the experiment. The aim of many studies on sequence learning is to demonstrate that participants are able to learn about repeated regularities in the sequence, and to determine whether this learning is accompanied by awareness or attributable to an implicit learning mechanism (e.g. Cleeremans & Jiménez, 1998; Jiménez, Méndez, & Cleeremans, 1996; Reber, 1989; Willingham, Nissen, & Bullemer, 1989). In such experiments, sequential effects are often regarded as variance to be controlled for or minimized on test (Jones, Curran, Mozer, & Wilder, 2013). The methods that researchers have employed to this end include devising an appropriate sequence to minimize sequential effects (e.g. avoiding first-order repetitions, Cleeremans & McClelland, 1991), or using a control group who are trained with a pseudorandom sequence containing no contingencies but with a trial order that would produce equivalent sequential effects (e.g. Anastasopoulou & Harvey, 1999; Jones & McLaren, 2009). The need to partial out sequential effects is a valid concern when attempting to measure sequence learning, since if sequential effects are merely performance effects, they may obscure or inflate evidence of learning (Vaquero, Jiménez, & Lupiáñez, 2006). Despite this, and the general treatment of sequential effects and sequence learning as separate phenomena in the literature, several researchers have suggested that sequential effects and sequence learning effects may result

* Corresponding author.
E-mail address: jessica.c.lee@sydney.edu.au (J.C. Lee).

http://dx.doi.org/10.1016/j.actpsy.2016.08.004

from the same learning mechanism (Audley, 1973; Laming, 1969; Soetens, Melis, & Notebaert, 2004).

Early studies on sequential effects were mostly confined to two-choice SRT tasks for the purpose of constraining the possible number of events (e.g. left and right) and transitions (repetitions and alternations of target location). For example, in a two-choice RT task (e.g. left and right responses) where the appearance of the target is randomly determined, participants are usually fastest to respond on trials where either repetitions or alternations of target location have occurred consecutively (e.g. Bertelson, 1961; Cho et al., 2002). This means that if a target had just appeared on the left 3 times, participants are usually faster to respond left than they are to respond right (i.e. LLLL would be faster than LLLR). Conversely, if participants have just experienced a series of alternations (left, right, left), they are faster at responding right than left (i.e. LRLR is faster than LRLL), but this facilitation is usually observed to be weaker than the equivalent effect for repetitions (e.g. Bertelson, 1961; Cho et al., 2002; Remington, 1969). These patterns of sequential effects have been attributed to participants' subjective expectancies (Soetens, Boer, & Hueting, 1985), which in this context refer to the predictions generated by some internal learning process. However, it is worth noting that these expectancies have been shown to be independent of the individual's explicit beliefs about impending events: recent work that has directly compared trends in choice RT and trends in explicit expectancy for the relevant events has found them to be widely divergent (Barrett & Livesey, 2010; Livesey & Costa, 2014; Lee Cheong Lem, Harris, & Livesey, 2015, see also Hale, 1967, and Hyman, 1953, for earlier informal observations of similar trends).

The internal learning process that leads to these subjective expectancies may be similar to the mechanisms that underlie sequence learning. Arguments in favor of a common mechanism include the fact that both sequence learning (e.g. Frensch & Miner, 1994) and sequential effects (e.g. Soetens, Boer, & Hueting, 1985) are highly sensitive to the length of the response-stimulus interval (RSI), practice affects sequence learning and sequential effects alike (Soetens et al., 2004), and the pattern of sequential effects is mirrored in electroencephalogram (EEG) studies investigating the P300 component, which is thought to code for prediction error (Squires, Wickens, Squires, & Donchin, 1976). These observations suggest that participants do form and update expectancies while responding to unstructured material, and thus the question of interest is what kind of mechanism leads to these expectancies. One possible answer is that sequential effects are a natural consequence of a rapid learning mechanism that is sensitive to short-term transient contingencies as well as long-term stable contingencies. In this way, sequential effects may represent a by-product of a highly adaptive ability to learn and change according to the statistics of a dynamic environment (Jones et al., 2013; Yu & Cohen, 2009).

Sequential effects models, however, have been largely developed separately of sequence learning models. Some models of sequential effects use simple associative architectures to represent the two-choice RT procedure, in combination with error-correction mechanisms. Despite variations between current models, there is some agreement that sequential effects in two-choice RT tasks can be explained by assuming that participants learn about the base rate of target locations (repetitions of specific target locations), and the frequencies of first-order transitions (repetitions and alternations of target location) (Jones et al., 2013; Wilder, Jones, & Mozer, 2009). Other successful attempts to model sequential effects have used detectors that track first-order contingencies to bias the system towards repetitions or alternations depending on trial history (Cho et al., 2002), or have omitted all hidden units and set up direct associations between representations of stimuli and responses (Gureckis & Love, 2010). These models of sequential effects provide good fits to empirical data and provide some indication of the statistics to which participants are sensitive. In contrast, models that have most successfully been applied to sequence learning incorporate similar learning principles with relatively complex model architecture, such as the augmented Serial Recurrent Network (SRN; Cleeremans & McClelland, 1991; Elman, 1990). If sequential effects are served by the same mechanisms as sequence learning, models like the augmented SRN, which is held to be the benchmark model of sequence learning (Beesley, Jones, & Shanks, 2012; Yeates, Jones, Wills, McLaren, & McLaren, 2013), should account for sequential effects to the same degree of success as they do for sequence learning involving complex deterministic and probabilistic transitions. The augmented SRN was purposefully modified from the original SRN (Elman, 1990) to account for short-term sequential effects (Cleeremans & McClelland, 1991), yet there has been relatively little reported work using the SRN to model sequential effects. Thus one of the aims of the current study was to test whether the augmented SRN could model sequential effects in addition to sequence learning effects in a novel three-choice RT task.

While there has been some research on sequential effects in choice-RT paradigms with more than two responses (Falmagne, 1965; Hyman, 1953; Schvaneveldt & Chase, 1969), these studies have mostly discussed the effects of repeating a single response location and have not fully examined other possible combinations of subsequences. One study that has investigated sequential effects in a three-choice SRT task used three different targets (geometric shapes), which could appear in the center of the screen, and participants responded by pressing the appropriate button using one finger on their dominant hand (Experiment 3, Gökaydin, Ma-Wyatt, Navarro, & Perfors, 2011). By comparing the sequential effects to an analogous procedure with only two possible targets (Experiment 2, Gökaydin et al., 2011), they concluded that adding an additional target caused participants to display sequential effects consistent with switching from tracking first-order statistics (repetitions and alternations of target location) to tracking base rate statistics (the relative frequency of each target). Their explanation was that introducing three possible responses increased task complexity, which in turn increased the number of possible first-order sequences that could be learned. Under these conditions they argued that participants reverted to learning about the base rates of each target, which was the simplest statistic to learn. This explanation implies a strategic and possibly intentional shift in the participant's learning strategy. It remains to be seen whether a sequence learning model like the SRN could account for changes in the number of target locations simply as a consequence of the changes in contingencies rather than a shift in attention to other event statistics.

In any case, Gökaydin et al.'s (2011) finding accords with the results of two-choice SRT tasks, where it is clear that first-order repetitions of target location (e.g. left-left-left) produce the most marked decreases in RT (e.g. Cho et al., 2002), and participants exclusively report noticing runs of target location when asked to explicitly look for a sequence before training (Experiment 3, Jones & McLaren, 2009). While there are important procedural differences in Gökaydin et al.'s (2011) task that reduce the generalizability to the majority of the two-choice RT literature (such as responding to the identity of the target rather than the location and only using one finger to respond), their findings highlight the importance of investigating sequential effects in different paradigms with more than two target locations in order to provide a more general account of sequential effects.

In the current study, we arranged three target locations on the edges of a computer screen (e.g. left-top-right) and prohibited repetitions of target location (e.g. top-top), to allow us to assess sequential effects concerning the repetition and alternation of the *direction* of target transitions, rather than target location (see Fig. 1). By prohibiting repetitions of target location, this task is similar to two-choice RT tasks in that on any given trial, there are only two possible events that can follow (a clockwise or anticlockwise transition). These spatial transitions add a novel and abstract quality to the SRT task, since direction of target movement (e.g. clockwise rotation) can summarize 3 different sets of contingencies (left-top, top-right, right-left). Using this paradigm, we assessed sequential effects by allowing an equal probability of clockwise or anticlockwise transitions (i.e. where the direction of motion on each trial is randomly determined, Experiment 1), and also assessed

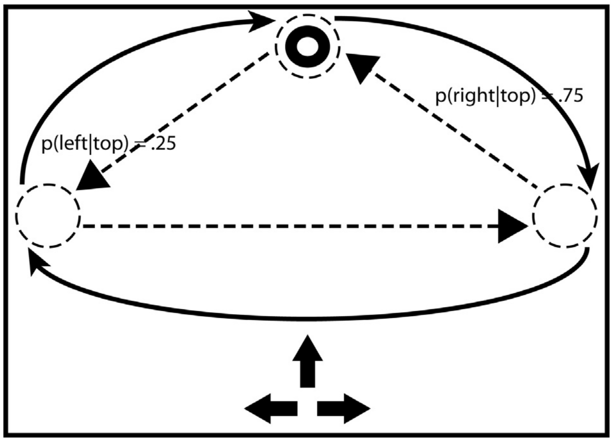

Fig. 1. Schematic diagram of the three-choice SRT task. The target could appear in either the left, top, or right position on screen (dotted circles, not seen in actual experiment) and participants had to respond by pressing the corresponding arrow key. The target could never appear in the same location twice in a row, which meant that the target would transition in a clockwise or anticlockwise direction on each trial (i.e. if the target appeared at the top, the next target location would either be left or right). In Experiment 1, the target direction was randomly determined so that there was a 50% chance of transitioning clockwise or anticlockwise on each trial. In Experiment 2, there was a cued direction of motion whereby the target would travel in a predominant direction 75% of the time (in this example, clockwise, as represented by the bold, curved lines), and in the miscued direction 25% of the time (anticlockwise, as represented by the straight, dotted lines).

probabilistic sequence learning by biasing the direction of motion in one direction (i.e. such that there is a 'cued' and 'miscued' direction of motion, Experiment 2).

The primary aim of the current study was to explore sequential effects in a three-choice SRT task and in particular, test whether the sequential effects regarding repetitions and alternations of target location also apply to other event statistics such as the *direction* of target transitions. While repetition effects clearly indicate learning of target location, alternation effects conflate the alternation of target *location* with the presence of target *motion* in terms of the transition from one target to the next. The suggestion is therefore that alternation effects may represent learning about target transitions, which is a higher-order statistic involving the transitional relation between one target and the next, and one to which both sequential effects and sequence learning may well be sensitive. We also tested whether these sequential effects and sequence learning in Experiment 2 can be modeled by the augmented SRN (Cleeremans & McClelland, 1991), since it is a popular model of sequence learning that should, in principle, be able to account for sequential effects. One of the advantages of the SRN is that it makes no assumptions about the statistics (e.g. base rates and repetition rates) that need to be learned and thus the statistics that the units code for are not 'hard-wired' into the model, removing the need to artificially change the event statistics that the model tracks when changing the number of target locations.

1. Experiment 1

The aim of Experiment 1 was to explore the sequential effects present in a three-choice SRT task that contained no response repetitions, such that all sequential effects would be based on sequences of transitions between target locations. Most studies examining sequential effects have used two-choice RT tasks, which means that our task is relatively novel despite the extensive literature on sequential effects (but see Gökaydin et al., 2011). Since neither direction of motion prevailed consistently, for any given target location, the other two target positions were equally likely to follow and therefore the target always moved in either a clockwise or anticlockwise direction. However, taking any three consecutive responses, the direction of motion itself would repeat if the three different target locations were shown consecutively in any order (in abstract terms, responses X, Y, then Z), whereas the direction of motion would reverse or alternate if the last response was the same as that occurring two presentations prior (that is, Z, Y, then Z again). A difference between these trial types can be thought of as second-order sequential effects (i.e. XYZ vs. ZYZ). A similar logic was applied to examine third-order sequential effects, that is, RT on the last of four consecutive responses that constitute three directional transitions, and fourth-order sequential effects, that is, RT on the last of five consecutive responses that constitute four directional transitions (see Fig. 2 and Table 1). A RSI of 500 ms was chosen because this delay between responses should be long enough to avoid response priming effects that dramatically alter sequential effects with short RSIs (<200 ms) (e.g. Vervaeck & Boer, 1980) but short enough for participants to retain a sense of directional transition from one target location to another.

1.1. Method

1.1.1. Participants and apparatus

Fifteen participants (11 female, *M* age = 26.87, *SD* = 7.90) who were either first year Psychology students at the University of Sydney or respondents to an online advertisement took part in the experiment. Students received course credit and respondents received payment (AUD$15/h) for their participation. All experiments were programmed using Psychophysics Toolbox for Matlab (Brainard, 1997; Pelli, 1997) and run on Apple Mac Mini desktop computers connected to 17 in. CRT monitors, refreshed at a rate of 85 Hz. A standard Apple keyboard and mouse were used, and testing was conducted in individual cubicles in groups of up to six. Participants gave informed consent and the study was approved by the University of Sydney Human Research Ethics Committee.

1.1.2. Procedure

Participants were told that the purpose of the task was to respond as quickly and as accurately as possible to a target (a magenta circle) that

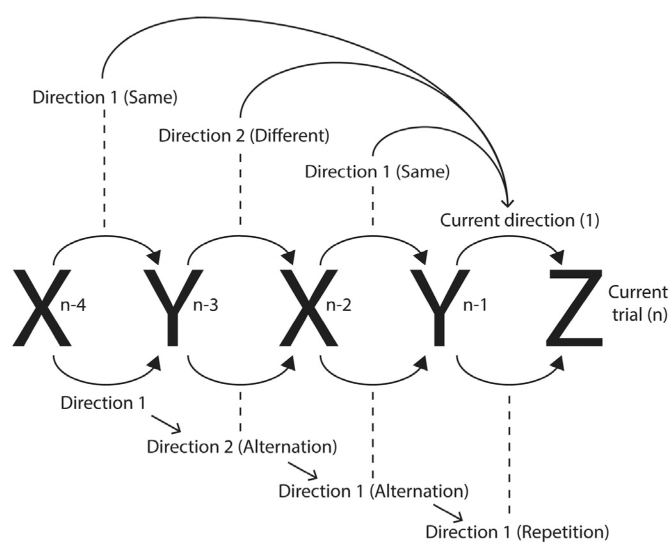

Fig. 2. Example of how a fourth-order subsequence was coded as a series of target locations (XYXYZ), a series of movements with reference to the direction of the first-order transition (SDS), and directional repetitions and alternations (AAR). X, Y, and Z can stand for any of the three target locations (left, top, right), with subsequences reading from left (past trials) to right (current trial, n), and therefore direction 1 and direction 2 can represent either clockwise or anticlockwise. Subsequences were entered into the ANOVA based on whether transitions at the nth level were in the same direction (S) or different direction (D) to the first-order transition (YZ, direction 1 in this example). Subsequences can also be conceived of as a series of repetitions (R) and alternations (A) of target direction, which are referenced from the direction of movement on the previous trial.

Table 1
Subsequences at fourth-, third-, and second-order level coded in three different ways.

Fourth order	Third order	Second order
SSS YZXYZ (RRR)	SS ZXYZ (RR)	S XYZ (R)
DSS XZXYZ (ARR)		
DDS ZYXYZ (RAR)	DS YXYZ (AR)	
SDS XYXYZ (AAR)		
DDD YXZYZ (RRA)	DD XZYZ (RA)	D ZYZ (A)
SDD ZXZYZ (ARA)		
SSD XYZYZ (RAA)	SD YZYZ (AA)	
DSD ZYZYZ (AAA)		

Note. The subsequences read from left (past trials) to right (current trial). X, Y and Z represent any one of the 3 target locations left, top, and right. R and A represent whether each subsequence consists of a repetition or alternation of direction, referenced from the previous trial. S and D represent whether the nth-order transition is the same (S) or different (D) direction to the 1st-order transition (YZ).

would appear in one of three positions on the screen. Participants were instructed to press the 'left' arrow key if the target appeared on the left, the 'up' arrow key if the target appeared at the top, and the 'right' arrow key if the target appeared on the right of the screen. Participants were not given any explicit instruction regarding the movement of the target (i.e., that it could not appear in the same location twice in a row) and were not encouraged to attend to the movement of the target. Participants were asked to use their non-dominant hand to respond during training. If participants used their left hand, they placed their ring finger on the left arrow key, their middle finger on the up arrow key, and their index finger on the right arrow key. The target stayed on screen until a response (correct or incorrect) was made and after a blank RSI of 500 ms, the next target appeared. After a short practice phase (48 trials), participants completed 720 trials where the location of the target had an equal chance of moving clockwise or anticlockwise and thus its location could not be predicted. Trials were randomized in blocks of 12, maintaining the 50/50 ratio of clockwise and anticlockwise transitions within each block. The experiment was completed in one continuous block without a break and lasted for approximately 15 min.

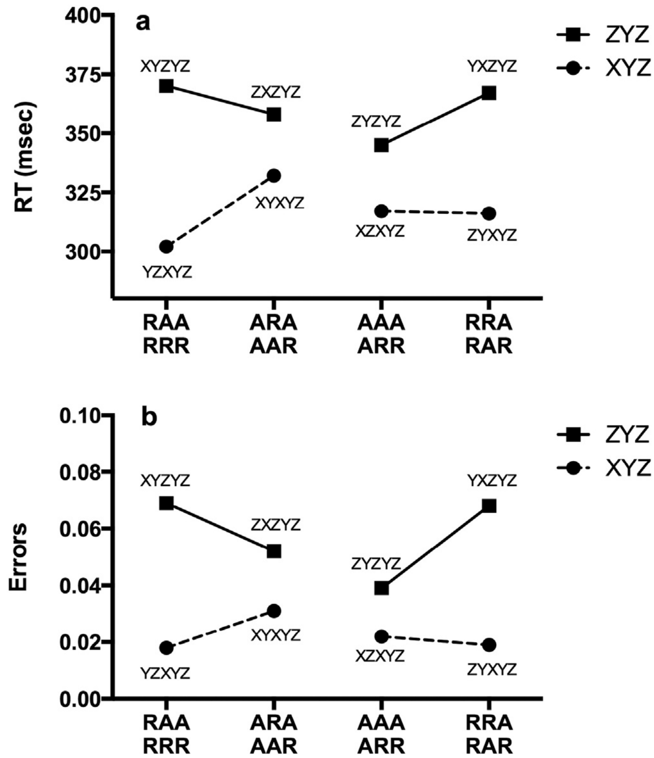

Fig. 3. RTs (a) and proportion of errors (b) for each fourth-order subsequence in Experiment 1. Subsequences are divided according to whether transitions at the second-order (XYZ vs. ZYZ, shown as separate lines), were the same or different direction to the first-order transition (YZ). Within each pair of connected data points, the left point has the same 3rd-order transition and the right has a different 3rd-order transition to the direction of the first-order transition (YZ). The pairs of connected data points on the left side of the figure have the same 4th-order transition, and the pairs on the right side of the figure have a different 4th-order transition to the first-order transition (YZ).

1.2. Results and discussion

All subsequent RT analyses refer to mean RTs for correct responses excluding any greater than one second[1] and Greenhouse-Geisser corrections were performed for violations of sphericity. Participants took on average 337 ms ($SD = 40.9$) to respond with 96% ($SD = 2.1$) accuracy. Henceforth X, Y, and Z will be used to describe the various subsequences with X, Y and Z representing any one of the 3 positions (left, top, right; see Fig. 1). This coding of subsequences is designed to capture the sequence of motion, but not the direction of movement. Thus, an XYZ subsequence could equally stand for either a left-top-right or left-right-top sequence. Reaction times and error data always represent performance on the final trial of each subsequence (Z). Trials were divided into subsequence type at the second-, third-, and fourth-order level (see Table 1). Within each level (n), trials were classified according to whether the transition at the nth level was the same (S) or different (D) from the direction of motion of the first-order transition (Y to Z, see Fig. 2 for an example of how a fourth-order subsequence was coded in this way), as well as whether the subsequence contained a series of alternations (A) or repetitions (R) of target direction. This yielded 2 different subsequences at the second-order level (R, A), 4 subsequences at the third-order level (RR, RA, AR, AA), and 8 subsequences at the fourth-order level (RRR, RRA, RAR, RAA, ARR, ARA, AAR, AAA).

Fig. 3 displays mean RTs and errors for the eight fourth-order subsequences, split according to whether the fourth-order transition was the same (left side of the figures) or different (right side of the figures) to the first-order transition. It is firstly apparent that the overall pattern of sequential effects is very similar for the RTs (top panels) and errors (bottom panels), and that there are very large differences between the two second-order subsequences (XYZ and ZYZ shown as separate lines), with performance on subsequences with a final repetition of motion (XYZ) faster than subsequences with a final alternation of motion (ZYZ). Within XYZ and ZYZ subsequences, the recent history of alternations and repetitions seemed to further impact performance.

To examine the pattern of sequential effects, a ($2 \times 2 \times 2$) ANOVA with fourth-order, third-order, and second-order as within-subjects factors was performed on mean RTs and errors for the fourth-order subsequences (Fig. 3). Note that the subsequences were entered into the ANOVA coded according to whether the nth transition was in the same or different direction to the first-order transition (see Table 1). For the RT data, there was a main effect of fourth-order, $F(1,14) = 6.40$, $p = 0.024$, $\eta_p^2 = 0.314$, and third-order, $F(1,14) = 5.31$, $p = 0.037$, $\eta_p^2 = 0.275$, and a very large main effect of second-order, $F(1,14) = 103.3$, $p < 0.001$, $\eta_p^2 = 0.881$. In the error data, there was also a main effect for second-order, $F(1,14) = 17.11$, $p = 0.001$, $\eta_p^2 = 0.550$, but no significant main effect of third- nor fourth-order, highest $F(1,14) = 2.68$, $p = 0.124$, $\eta_p^2 = 161$. It is clear that the strongest main effect in both RT and errors was at the second-order level, specifically comparing the XYZ subsequence (RT: $M = 317$ ms, $SD = 39.5$, errors: $M = 0.023$, $SD = 0.013$) to the ZYZ subsequences (RT: $M = 363$ ms, $SD = 44.1$, errors: $M = 0.061$, $SD = 0.035$).[2] Participants were on average 46 ms faster and also made on average 3.8% fewer errors when the target travelled in a consistent direction on 2 consecutive transitions (those ending in XYZ), compared to when the target appeared to

[1] This resulted in discarding of 4.3% of trials due to errors or RTs > 1 s.

[2] The second-order sequential effects were previously reported in the Proceedings of the Cognitive Science Society (Lee & Livesey, 2013).

alternate directions (those ending in ZYZ). The main effects of third- and fourth-order in the RT data show that participants were 10 ms faster to respond when the third-order transition was the same direction as the first (subsequences of the form _S_ were faster than _D_, i.e. in Fig. 3, of the points connected by lines the left points are lower than the right points), but 4 ms slower to respond when the fourth-order transition was the same direction as the first (subsequences of the form S_ _ were slower than D_ _, i.e. in Fig. 3, the left hand side of the figures is higher than the right hand side). Note however, that these main effects are qualified by the significant interactions discussed below.

In RTs, there was a significant third-order x second-order interaction, $F(1,14) = 7.12$, $p = 0.018$, $\eta_p^2 = 0.337$, and significant fourth-order x second-order interaction, $F(1,14) = 5.31$, $p = 0.037$, $\eta_p^2 = 0.275$. A significant 3-way interaction between fourth-, third-, and second-order factors was found in both RTs, $F(1,14) = 31.34$, $p < 0.001$, $\eta_p^2 = 0.691$, and accuracy, $F(1,14) = 16.13$, $p = 0.001$, $\eta_p^2 = 0.535$ (see Fig. 3). The easiest way to interpret the three-way interaction is by conceptualizing the subsequences as a series of directional repetitions and alternations (see Table 1). If we examine the 4 fourth-order subsequences where the fourth-order transition was consistent with the first-order transition (left side of the figures), it is clear that RT and accuracy were influenced primarily by second-order differences. That is, whether the last trial in the subsequence contained a repetition (R) or alternation (A). However, within the XYZ subsequences, responding was facilitated when the subsequence contained several repetitions in a row (RRR was easier to respond to than AAR), and within the ZYZ subsequences, responding was facilitated when the subsequence contained a repetition before the last alternation (ARA), compared to when there were 2 alternations to respond after (RAA). This pattern seems to reverse for the examination of those subsequences where the fourth-order transition is inconsistent with the first-order transition (right side of the figures). For the XYZ subsequences, there seemed to be little difference between whether the third-order transition contained a repetition (ARR) or alternation (RAR), as responding seemed to be generally facilitated by the repetition of a direction of motion on the last trial of the subsequence. On the other hand, for the ZYZ subsequences, responding was both faster and more accurate when the subsequence contained a series of alternations in a row (AAA) than when it contained a series of repetitions and then a final alternation (RRA). The pattern of data in Experiment 1 can be summarized in the following way: general facilitation in responding occurred when the target moved in the same direction a few times in a row (i.e. there was a repetition of a direction of motion), and responding was hindered when the direction alternated, except when the direction alternated several times (i.e. ZYZYZ).

It is clear from this experiment that higher-order sequential effects exist in this task, and while some interactions between second-, third- and fourth-order levels of subsequences were significant in this experiment, by far the most substantial difference was at the second-order level between the XYZ and ZYZ subsequences. The biggest determinant for whether responding in this task was facilitated was whether the previous direction of motion was consistent with the current direction of motion. It appears that sequences of trials in which the target changed direction led to slower responses (the fastest ZYZ subsequence, ZYZYZ, was still numerically slower than the slowest XYZ subsequence, XYXYZ, see Fig. 3).

2. Experiment 2

In Experiment 2, contingencies were added to the task such that the target would appear to be moving in one direction 75% of the time during training, and a transfer phase was also added where these contingencies were removed. This transfer phase allows us to test whether participants had learnt about the contingencies, with learning evident if participants are faster to respond on trials where the target moved in the previously cued direction of motion (cued trials) than the previously miscued direction (miscued trials). To equate the analysis between Experiment 1 and Experiment 2, we focused on analyzing the sequential and cueing effects during the transfer phase only, such that any sequential effects shown would be under conditions where there was no predictive information.

2.1. Method

2.1.1. Participants

All fifteen participants (9 female, M age $= 25.47$, $SD = 7.22$) in Experiment 2 were respondents to an online advertisement and were paid AUD $15/h for their participation.

2.1.2. Procedure

The procedure was identical to Experiment 1 except for the following changes. After completing a short practice phase (48 trials) with no contingencies, participants responded to 720 trials where the target moved in a prevailing direction of motion on 75% of trials (which was randomly chosen to be clockwise or anticlockwise for each participant) and 360 trials where there were no contingencies (there was no prevailing direction of motion). For the initial 720 trials with prevailing direction of motion, trials were randomized in blocks of 12 trials maintaining the 75% cued and 25% miscued ratio of contingencies within each block. Participants continued to use the same response keys and hand to respond and there was no break between the training and transfer phase, such that there was nothing to mark the transition into a separate phase for participants. The instructions given to participants were exactly the same as Experiment 1, meaning that participants were not informed that there was a bias in the direction of motion.

2.2. Results and discussion

The data were analyzed in a similar way to Experiment 1, with cueing added as a within-subjects factor.[3] A $(2 \times 2 \times 2 \times 2)$ within-subjects ANOVA with cueing, fourth-, third-, and second-order as factors was run on RTs (Fig. 4a) and errors (Fig. 4b) in the transfer phase. We purposefully restricted our analyses to the transfer phase to equate the assessment of sequential effects, as much as possible, to that conducted on the data from Experiment 1, where there was an absence of contingencies (see Supplementary Materials for the results of the analyses on the training phase). In RTs, there was a significant main effect of cueing, $F(1,14) = 45.96$, $p < 0.001$, $\eta_p^2 = 0.767$, with faster responses for cued trials indicating that participants expected transitions to occur in the cued direction during the transfer phase. There was also a main effect of second-order, $F(1,14) = 42.02$, $p < 0.001$, $\eta_p^2 = 0.750$, and a significant interaction between fourth-, third-, and second-order, $F(1,14) = 49.31$, $p < 0.001$, $\eta_p^2 = 0.779$. In errors, there was a main effect of second-order, $F(1,14) = 9.25$, $p = 0.009$, $\eta_p^2 = 0.398$, a significant interaction between fourth- and third-order, $F(1,14) = 12.81$, $p = 0.003$, $\eta_p^2 = 0.478$, and also a significant 3-way interaction between fourth-, third-, and second-order, $F(1,14) = 13.43$, $p = 0.003$, $\eta_p^2 = 0.490$. This broadly replicates the sequential effects found in Experiment 1, where participants showed a very large difference in responding between XYZ and ZYZ subsequences, and produced both repetition and alternation effects that explain the 3-way interaction. Interestingly, while participants showed very strong cueing effects overall, the 4-way interaction was not significant in either RTs nor errors, $F < 1$, and nor were any other interactions significant, $Fs(1,14) \leq 2.06$, $ps \geq 0.174$, $\eta_p^2 \leq 0.128$, suggesting that the pattern of sequential effects shown on cued and miscued trials was very similar (see Fig. 4).

[3] 3.2% of trials were discarded from the RT analysis due to errors or being >1 s.

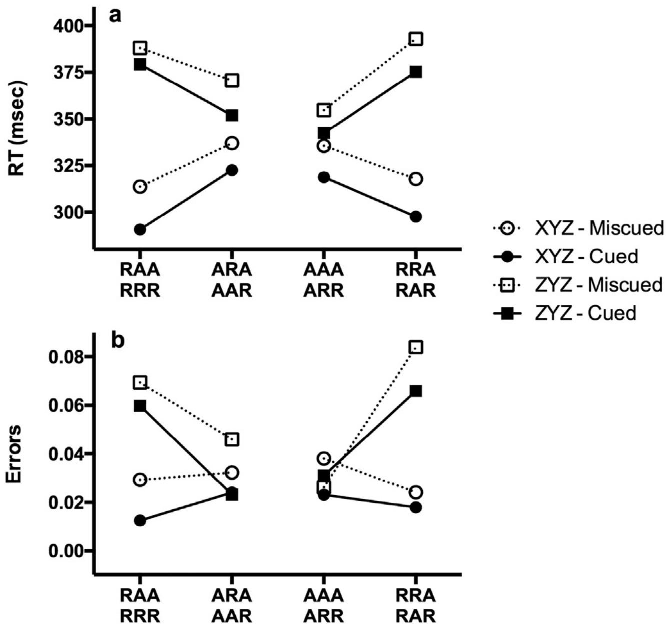

Fig. 4. RTs (a) and proportion of errors (b) for cued and miscued trials for each fourth-order subsequence in the transfer phase of Experiment 2. Subsequences are divided according to whether transitions at the second-order (XYZ vs. ZYZ, shown as separate lines), third-order (left vs. right points connected by lines), and fourth-order (left vs. right side of the figures) level were the same or different direction to the first-order transition (YZ). Cued and miscued trials are shown as separate lines.

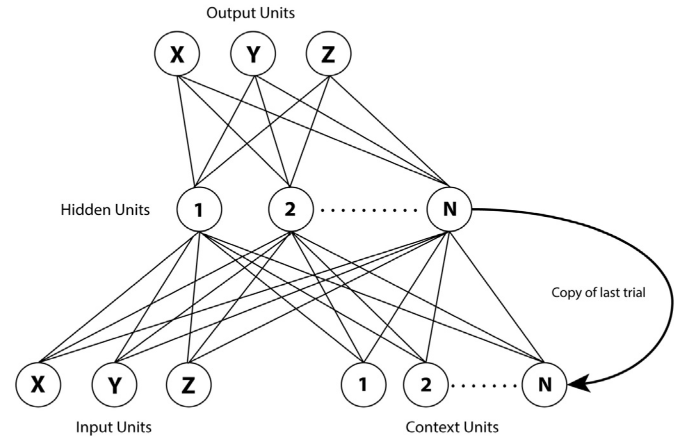

Fig. 5. Model architecture for the augmented SRN with three target locations and N hidden units. The three input and output units (X, Y, Z) correspond to the three target locations (left, top, and right of screen). The network is recurrent in that the activations in the hidden layer at time t are copied back to a set of context units which feed into the next trial at time $t + 1$.

3. General discussion

Using a novel three-choice RT task where the target locations were arranged on the left, top and right of a computer screen, we found a robust pattern of sequential effects in the absence of a cued direction of motion in Experiment 1, and for cued and miscued trials in a transfer phase following training with a biased direction of motion. Since the target never appeared in the same location on successive trials, the target location was constrained to always move in one of two directions: clockwise or anticlockwise. When the contingencies were biased in one direction 75% of the time, participants appeared to learn this probabilistic sequence by showing a cueing effect once the contingencies were removed in the transfer phase of Experiment 2. While RTs in the transfer phase were generally faster for cued than for miscued trials, the pattern of sequential effects did not appear to be different between cued and miscued trials, suggesting an additive effect of short- and long-term contingencies on RTs. Participants responded fastest to subsequences containing repetitions of target direction (YZXYZ trials), similar to consistent repetition advantages observed in the two-choice RT literature (e.g. Bertelson, 1961; Hyman, 1953). Interestingly, for the ZYZ subsequences, we found that the subsequence that was responded to most rapidly was the one where the target direction alternated consistently (ZYZYZ), again similar to the findings using two-choice RT tasks (e.g. Cho et al., 2002). It thus appears that, behaviorally speaking, participants anticipate runs of alternations and repetitions to continue not just for target location, but also for the direction of target motion, at least in a task where target location repetitions are prohibited.

Since our aim was to see whether a model of sequence learning could also be used to model sequential effects, we chose the augmented SRN (for further details see Cleeremans & McClelland, 1991) to simulate our results. The SRN (Elman, 1990) is a connectionist model (Rumelhart, Hinton, & Williams, 1986), which uses a simple error-correction learning algorithm to simulate human behavior. The SRN is composed of a layer of input units which code for the presence of external stimuli, an output layer whose activation reflects the model's prediction about the next item in the sequence, and a hidden layer that connects the input layer to the output layer (see Fig. 5). The SRN builds a representation of the sequence of events across time and allows for previous items in the sequence to affect the prediction regarding the next element. It achieves this by copying the pattern of activation across the hidden units on trial t-1 to a set of context units, which provide additional input to the hidden units on trial t. This recurrent nature of the network means that previous trials influence the current trial via the hidden units.

We used a network architecture with three input and three output units, representing the three target locations. The activation in both the hidden layer and input layer were determined using the logistic function, and the activation in the output units represented the network's prediction on a given trial. The predicted response efficiency generated by the model was calculated by taking Luce's ratio rule (Luce, 1959) of the activations of the outputs, that is, the activation of the correct target output divided by the sum of the activations of all output units. Since RT is inversely proportional to prediction strength and response efficiency, model output was then calculated by subtracting this response efficiency estimate from 1. The weights were modified according to a back-propagation algorithm, which acts to reduce error for each prediction on a trial-by-trial basis (Rumelhart et al., 1986).

The augmented SRN was originally modified from the SRN to account for what Cleeremans and McClelland (1991) referred to as short-term priming. In their experiments, they observed that response times differed across a variety of sequences of transitions, irrespective of the structured nature of these sequences (e.g. RTs for QXQX were faster than VXQX despite both sequences being equally permissible). They suggested that these patterns in responding could be explained by priming of sequential pairs of responses (e.g. QX primes activation of the subsequent QX). In addition to a set of slow weights that were used to represent relatively permanent long-term sequence learning, Cleeremans and McClelland included in their model a set of ‘fast’ weights with a fast decay rate that would allow the model to produce short-term priming effects. This effectively means that each association (e.g. Q-X) has a short-term and long-term component (or fast and slow weight), which combine to determine output in the model. The short-term component produces a large increment in the connection weight but with a rapid decay rate (half-life of one time step), while the long-term component produces a smaller increment with a much slower decay rate. This allows the augmented SRN to predict fast responding for subsequences where a pair of stimuli are repeated (sequential effects), as well as long-term learning effects (sequence learning).

To see whether the augmented SRN could simulate the results, we ran a simulated annealing procedure (Kirkpatrick, Gelatt, & Vecchi, 1983) to find the best fitting parameters (slow learning rate within a

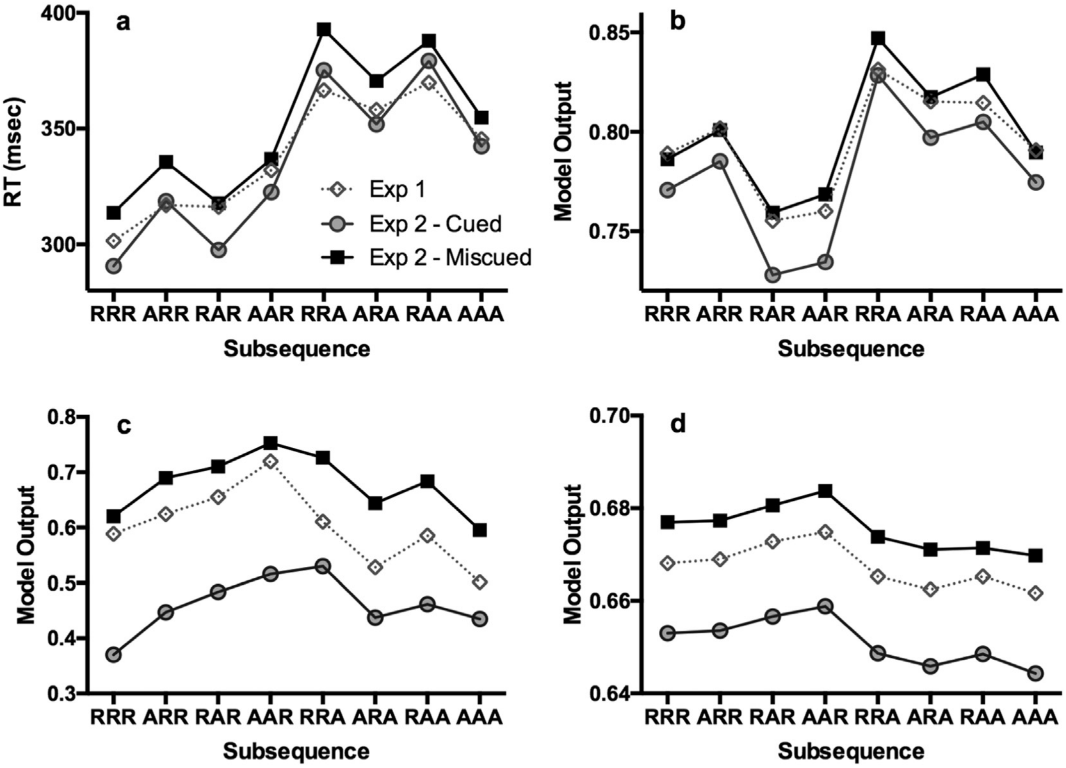

Fig. 6. RTs for each fourth-order subsequence from Experiments 1 and 2 (a) and corresponding model output using 3 variations of the SRN (b-d). B) Model output from the augmented SRN using a fast learning rate of 1.00, a slow learning rate of 0.029, and 45 hidden units. C) Model output from the original SRN with a single learning rate of 1.31 and 24 hidden units. D) Model output from the augmented SRN with the context layer removed using a slow learning rate of 0.089 and fast learning rate of 0.574. Note the different scales on each figure.

range of 0.01 to 1, fast learning rate within a range of 0.1 to 2.5, and number of hidden units within a range of 1 to 50) to model the ordinal fit of the fourth-order subsequence RTs (using Spearman's rho) from Experiments 1 and 2 (Fig. 6a).[4] This procedure gave us a slow learning rate of 0.03, a fast learning rate of 1.0, and 45 hidden units (see Table 2). Note that the learning rates are strength parameters governing the adjustments in weights and should not be interpreted in absolute terms (i.e. the fast learning rate of 1 does not mean that the most recent trials contributed 100% to the weight changes). Rather, a small learning rate corresponds to small weight changes, and a large learning rate corresponds to large weight changes. The network was then fed the trial sequence as generated for the actual participants for both Experiment 1 and 2. This meant that the practice trials, training and transfer phases were simulated with the same number of trials. Using these parameters, the results of a simulation involving 20 participants for each of Experiments 1 and 2 is shown in Fig. 6b, along with the empirical RTs obtained from both Experiments in Fig. 6a.

The first observation to highlight is that the general order of the eight subsequences is mostly preserved, and consistent with Experiment 2 there are roughly equivalent cueing effects across each subsequence. However, although the model predictions are largely correct at the second order level, the ordinal fit at the fourth-order level is certainly not perfect. The augmented SRN has successfully captured the overall advantage for XYZ subsequences (ending in R) compared to ZYZ subsequences (ending in A) (i.e. the left half of the figure is lower than the right half), however the variation within each of the four XYZ and ZYZ subsequences was not entirely replicated. In particular, when comparing the model output to the empirical RTs, the SRN predicts a relatively fast response to XYXYZ (AAR) subsequences, and a relatively slow response to YZXYZ (RRR) subsequences. Interestingly, both slow responses to AAR and fast responses to RRR are consistent with participants learning about sequences of repetitions and alternations of target *direction*, something which the SRN cannot explicitly represent. In particular, the SRN's poor prediction of fast RTs for YZXYZ (RRR) subsequences makes sense, since on the current trial (Z), both X and Y appear more recently in the trial sequence and should be more strongly primed than Z. Thus, there is no reason why responses should be faster for Z over X and Y, unless the transition from *Z*-X and then from X-Y are priming the transition from Y-Z. The SRN's over-prediction of fast responses for XYXYZ (AAR) is more mysterious, since the Y-Z transition occurs more recently on YZXYZ (RRR) trials and yet the SRN predicts slower responses on these trials. While it is obvious that the augmented SRN fails to capture these aspects of the data, overall, it does an adequate job of simulating the results using only local representation.

The augmented SRN is quite different from the models traditionally used within the sequential effects literature, and while it is not our aim to show that the SRN is superior to existing models of sequential effects, it may well be the case that a simpler model might be able to simulate our results. Still, it is possible to demonstrate that specific properties of the augmented SRN are critical in allowing it to generate the empirical data, by repeating the above procedure with modified versions of the SRN (see Table 2 for a comparison of the optimal parameters). The first version was the SRN in its original form without a set of fast weights

Table 2
Best fitting parameters found using a simulated annealing procedure for each version of the SRN.

SRN version	Slow learning rate	Fast learning rate	N hidden
Augmented	0.029	1.00	45
Original	1.31	0	24
Augmented minus hidden	0.089	0.574	0

[4] Simulated annealing is a fast parameter optimization procedure that is less susceptible to problems with local minima than conventional path searches. We used the simulannealbnd function that is available as part of the Global Optimization Toolbox for MatLab to run the parameter search.

(i.e., Elman, 1990). The model output using the best fitting parameters (learning rate of 1.31, 24 hidden units) can be seen in Fig. 6c. It is apparent that the model is certainly able to predict an overall cueing effect, but it is large in magnitude relative to the size of the sequential effects. It is also clear that the model fails to replicate the pattern of sequential effects, with the most glaring discrepancy being that of the second-order sequential effects. Without the fast learning rate, the model incorrectly predicts that ZYZ subsequences (ending in A) will be faster than XYZ subsequences (ending in R) (i.e. the left side of the figure is higher than the right side of the figure).

The second version of the SRN was one in which the hidden layer was removed such that there was no internal transformation of event representation (input units were directly connected with outputs via modifiable slow and fast weights). Fig. 6d shows the model output with the best fitting parameters (slow learning rate of 0.089 and fast learning rate of 0.574). Again, the model is able to predict a large cueing effect but the pattern of sequential effects is not reproduced. The issue again concerns the incorrect prediction of the direction of the second-order sequential effects, with the model predicting ZYZ subsequences to be faster than XYZ subsequences. Both modified versions of the SRN also fail to capture the variation amongst each of the XYZ and ZYZ subsequences. Thus, it is clear that both the fast learning rate and the hidden layer in the augmented SRN are critical for the prediction of our obtained results.

It should be noted that the optimal parameters we found are also quite different in comparison to other studies that have used the augmented SRN (e.g. Cleeremans & McClelland, 1991; Yeates et al., 2013). For example, our fast learning rate is very high and our slow learning rate is very low in comparison to that used by Cleeremans and McClelland (1991), who used 0.15 as their slow learning rate and 0.2 as their fast learning rate. This may be because relative to the size of the sequential effects, the size of the cueing effect was quite small in magnitude in our experiments (though consistent across subsequences), and the statistic we used to fit the SRN took this into account. Within the parameter space that we considered, the SRN generally overestimates the size of the cueing effect relative to the size of sequential effects. Having a very fast learning rate for the fast decaying weights serves to increase the magnitude of the sequential effects while keeping the more stable long-term cueing effect fairly modest.

By focusing on the augmented SRN, we seek to illustrate that sequential effects and sequence learning effects should be investigated using the same theoretic approach, especially since they emerge from the same response paradigms and on the same measures, and their relative magnitudes could be informative for further model development. As discussed above, an advantage of the SRN is that it makes no assumptions about the types of statistics that it needs to monitor, since representations of repetitions, alternations and target direction are not explicitly 'hard-wired' into the model. Indeed, the augmented SRN managed to model our results reasonably well without needing to add anything in addition to its representations of each of the three target locations. This stands in contrast to sequential effects models where the coding of each unit needs to be determined and implemented explicitly (e.g. Cho et al., 2002; Jones et al., 2013). However, this feature of the SRN also makes it far less transparent. For instance, although the model proposed by Jones et al. (2013) requires a mechanism for the explicitly coded tracking of event statistics, it is clear what the model is learning in order to produce sequential effects. It is much less clear how the sequential representations developed within the architecture of the SRN allow it to arrive at the predictions illustrated in Fig. 6b. The SRN has been noted for being sensitive to non-associative structure and almost symbolic in its computations despite its connectionist architecture due to the representational power of its hidden layer (Gureckis & Love, 2010). While the structure in our task need not be described in non-associative terms, the large number of hidden units needed to model our results does suggest that a complex level of representation was needed in order to simulate our results. Thus while the SRN was capable of reproducing our results, it does not provide a precise answer as to the content of learning.

Despite the nature of the task making the direction of the target salient in general, it is unclear whether participants learned something about the target direction in particular, or whether they simply learned about contingencies between specific target locations. Even in Experiment 2 when the direction of motion was biased in the cued direction (clockwise, for example) and participants showed a reliable cueing effect, participants could be displaying this cueing effect because they have an expectation that the target travels clockwise most of the time, or because they have an expectation based on the individual contingencies (i.e. they learned that top follows left, right follows top and left follows right). Knowing that the target moves in the cued direction 75% of the time entails knowledge about the individual probabilistic contingencies as well, but it is possible to learn the contingencies in the absence of the abstract relationship concerning movement in a clockwise or anticlockwise direction.[5] Thus, the presence of cueing effects in Experiment 2 does not necessarily indicate that participants learned something more abstract than the individual contingencies, nor whether this knowledge is able to be expressed explicitly. Future work could explore the nature of this knowledge by implementing various transfer phases where, for example, the target locations are changed but the direction of motion kept consistent.

While we did not manipulate task features such as RSI, the mapping between target location and responses, or placement of the target locations around the screen, the literature on two-choice SRT suggests that task features such as these will almost certainly influence the observable pattern of sequential effects. In particular, our task did not allow for first-order repetitions, which makes it quite different from the three-choice RT task in Gökaydin et al.'s (2011) study. Since Gökaydin et al. found that when three different targets were used, participants primarily showed sequential effects consistent with learning about the base rates of stimuli, it would be interesting to see how allowing first-order repetitions would change the pattern of sequential effects shown here. Gökaydin et al. also deliberately avoided using spatial mapping from response keys to target location in order to reduce sequential effects generated by the target placement. It would thus be interesting to see whether adapting our probabilistic sequence into a task similar to theirs where different targets appear in the same location on screen would generate the same sequential effects and to the same degree. In particular, the large advantage in RT for XYZ subsequences over ZYZ subsequences may be reversed since alternations of targets (ZYZ) may be more salient or easier to respond to than sequences of all different targets (XYZ) when participants have to remember which key to press for different targets. Whether the SRN would be able to cope with changing these task features when it can only represent three abstract target locations or identities, remains to be seen.

In conclusion, the current study demonstrates that similar sequential effects to that found with repetition and alternation of target location in the two-choice SRT literature can also be found in other event statistics in three-choice SRT such as repetition and alternation of the direction of target motion. We introduced a novel three-choice SRT task with a relational property (direction of target movement), which we then used to bias the contingencies. We found reliable sequential effects in both Experiment 1 where there was no predominant direction of motion, and Experiment 2 where there was a cued direction of motion. Experiment 2 also showed that the pattern of sequential effects did not

[5] It is possible that the use of the arrow keys would have made participants more sensitive to detecting motion in the sequenced transitions, since these keys are typically used to move cursors and to make similar actions with objects on computer screens. Note however that the arrow keys do not convey anything about *rotational* motion and thus would not be strongly suggestive of the dominant direction of motion that was implemented in Experiment 2, even though their use may be suggestive of target movement in a general sense. Any specific direction of motion conferred to the participants by these keys would have been as likely to hinder, as enhance, the learning of motion in the sequences, due to the strict counterbalancing of the task elements.

differ between cued and miscued trials once those contingencies were removed. Participants in both experiments were generally faster on XYZ trials where the direction of motion was repeated compared to ZYZ trials where the direction of motion alternated, but within each of these second-order subsequences, participants were fastest to respond on runs of repetitions or alternations (respectively) of target direction. We successfully simulate these cueing and sequential effects with reasonable accuracy using the augmented SRN (Cleeremans & McClelland, 1991) provided the model contained a relatively large number of hidden units, a high learning rate for fast weights and a low learning rate for slow weights. While the model failed to capture the more subtle effects in the 4th order subsequences, the overall fit provided by the augmented SRN was promising and suggests a common basis for sequence learning and sequential effects.

Acknowledgements

This research was supported by an Australian Research Council Discovery grant DP120102190 to EJL.

The authors declare that no competing interests exist.

Experiment 1 was previous reported in the Proceedings of the Cognitive Science Society (Lee & Livesey, 2013).

The authors would like to thank Hilary Don for helpful comments on an earlier draft, and Dinis Gökaydin for valuable advice and discussion.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.actpsy.2016.08.004.

References

Anastasopoulou, T., & Harvey, N. (1999). Assessing sequential knowledge through performance measures: The influence of short-term sequential effects. *The Quarterly Journal of Experimental Psychology*, *52A*, 423–448.

Audley, R. (1973). Some observations on theories of choice reaction time: Tutorial review. *Attention and performance IV* (pp. 509–545).

Barrett, L. C., & Livesey, E. J. (2010). Dissociations between expectancy and performance in simple and two-choice reaction-time tasks: A test of associative and nonassociative explanations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 864–877.

Beesley, T., Jones, F. W., & Shanks, D. R. (2012). Out of control: An associative account of congruency effects in sequence learning. *Consciousness and Cognition*, *21*, 413–421.

Bertelson, P. (1961). Sequential redundancy and speed in a serial two-choice responding task. *The Quarterly Journal of Experimental Psychology*, *13*, 90–102.

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

Burns, B. D., & Corpus, B. (2004). Randomness and induction from streaks: "Gambler's fallacy" vs. "Hot hand". *Psychonomic Bulletin and Review*, *11*, 179–184.

Cho, R. Y., Nystrom, L. E., Brown, E. T., Jones, A. D., Braver, T. S., Holmes, P. J., & Cohen, J. D. (2002). Mechanisms underlying dependencies of performance on stimulus history in a two-alternative forced-choice task. *Cognitive, Affective, & Behavioral Neuroscience*, *2*, 283–299.

Cleeremans, A., & Jiménez, L. (1998). Implicit sequence learning: The truth is in the details. In M., & P. (Eds.), *Handbook of implicit learning* (pp. 323–364). Sage: Newbury Park, CA.

Cleeremans, A., & McClelland, J. L. (1991). Learning the structure of event sequences. *Journal of Experimental Psychology: General*, *120*, 235–253.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, *14*, 179–211.

Falmagne, J. C. (1965). Stochastic models for choice reaction time with applications to experimental results. *Journal of Mathematical Psychology*, *2*, 77–124.

Frensch, P. A., & Miner, C. S. (1994). Effects of presentation rate and individual differences in short-term memory capacity on an indirect measure of serial learning. *Memory & Cognition*, *22*, 95–110.

Gökaydin, D., Ma-Wyatt, A., Navarro, D., & Perfors, A. (2011). Humans use different statistics for sequence analysis depending on the task. In L., C., & T. (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Society* (pp. 543–548). Austin, TX: Cognitive Science Society.

Gureckis, T. M., & Love, B. C. (2010). Direct associations or internal transformations? Exploring mechanisms underlying sequential learning behavior. *Cognitive Science*, *34*, 10–50.

Hale, D. J. (1967). Sequential effects in two-choice serial reaction task. *The Quarterly Journal of Experimental Psychology*, *19*, 133–141.

Hyman, R. (1953). Stimulus information as a determinant of reaction time. *Journal of Experimental Psychology*, *45*, 188–196.

Jarvik, M. E. (1951). Probability learning and a negative recency effect in the serial anticipation of alternative symbols. *Journal of Experimental Psychology*, *41*, 291–297.

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.

Jones, F. W., & McLaren, I. P. L. (2009). Human sequence learning under incidental and intentional conditions. *Journal of Experimental Psychology: Animal Behavior Processes*, *35*, 538–553.

Jones, M., & Sieck, W. R. (2003). Learning myopia: An adaptive recency effect in category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 626–640.

Jones, M., Curran, T., Mozer, M. C., & Wilder, M. H. (2013). Sequential effects in response time reveal learning mechanisms and event representations. *Psychological Review*, *120*, 628–666.

Jones, M., Love, B. C., & Maddox, W. T. (2006). Recency effects as a window to generalization: Separating decisional and perceptual sequential effects in category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 316–332.

Kirkpatrick, S., Gelatt, C. D., & Vecchi, M. P. (1983). Optimization by simulated annealing. *Science*, *220*, 671–680.

Laming, D. R. J. (1969). Subjective probability in choice-reaction experiments. *Journal of Mathematical Psychology*, *6*, 81–120.

Lee, J. C., & Livesey, E. J. (2013). Effects of explicit abstract knowledge and simple associations in sequence learning. In M., M., N., & I. (Eds.), *Proceedings of the 35th Annual Convention of the Cognitive Science Society* (pp. 2843–2848). Austin, TX: Cognitive Science Society.

Lee Cheong Lem, V. A., Harris, J. A., & Livesey, E. J. (2015). Testing the limits of the Perruchet effect in choice response time tasks. *Journal of Experimental Psychology: Animal Learning and Cognition*, *41*, 385–394.

Link, B. V., Kos, B., Wager, T. D., & Mozer, M. C. (2011). Past experience influences judgment of pain: Prediction of sequential dependencies. In L., C., & T. F. (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 1248–1253). Austin, TX: Cognitive Science Society.

Livesey, E. J., & Costa, D. S. J. (2014). Automaticity and conscious control in single and choice response time versions of the Perruchet effect. *The Quarterly Journal of Experimental Psychology*, *67*, 646–664.

Luce, R. D. (1959). *Individual choice behavior: A theoretical analysis.* New York: Wiley.

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, *19*, 1–32.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.

Perruchet, P. (1985). A pitfall for the expectancy theory of human eyelid conditioning. *Pavlovian Journal of Biological Science*, *20*, 163–170.

Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General*, *118*, 219–235.

Remington, R. J. (1969). Analysis of sequential effects in choice reaction times. *Journal of Experimental Psychology*, *82*, 250–257.

Rumelhart, D. E., Hinton, G., & Williams, R. J. (1986). Learning internal representations by error propagation. In D. E., & J. L. (Eds.), *Parallel distributed processing: I. Foundations* (pp. 318–362). Cambridge, MA: MIT Press.

Schvaneveldt, R. W., & Chase, W. G. (1969). Sequential effects in choice reaction time. *Journal of Experimental Psychology*, *80*, 1–8.

Soetens, E., Boer, L. C., & Hueting, J. E. (1985). Expectancy or automatic facilitation? Separating sequential effects in two-choice reaction time. *Journal of Experimental Psychology: Human Perception and Performance*, *11*, 598–616.

Soetens, E., Melis, A., & Notebaert, W. (2004). Sequence learning and sequential effects. *Psychological Research*, *69*, 124–137.

Squires, K. C., Wickens, C., Squires, N. K., & Donchin, E. (1976). The effect of stimulus sequence on the waveform of the cortical event-related potential. *Science*, *193*, 1142–1146.

Vaquero, J. M. M., Jiménez, L., & Lupiáñez, J. (2006). The problem of reversals in assessing implicit sequence learning with serial reaction time tasks. *Experimental Brain Research*, *175*, 97–109.

Vervaeck, K. R., & Boer, L. C. (1980). Sequential effects in two-choice reaction time: Subjective expectancy and automatic after-effect at short response-stimulus intervals. *Acta Psychologica*, *44*, 175–190.

Wilder, M., Jones, M., & Mozer, M. C. (2009). Sequential effects reflect parallel learning of multiple environmental regularities. In Y., D., J., C. K. I., & A. (Eds.), *Advances in neural information processing systems 23* (pp. 2053–2061). La Jolla, CA: NIPS Foundation.

Willingham, D. B., Nissen, M. J., & Bullemer, P. (1989). On the development of procedural knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 1047–1060.

Yeates, F., Jones, F. W., Wills, A. J., McLaren, R. P., & McLaren, I. P. L. (2013). Modeling human sequence learning under incidental conditions. *Journal of Experimental Psychology: Animal Behavior Processes*, *39*, 166–173.

Yu, A., & Cohen, J. (2009). Sequential effects: Superstition or rational behavior? In Y., D., J., C. K. I., & A. (Eds.), *Advances in neural information processing systems 22* (pp. 1873–1880). La Jolla, CA: NIPS Foundation.