UNIVERSITY OF CAMBRIDGE

Impact of Task Structure on Strategies in Statistical Learning

MONICA ANNA TANG GATES

Supervisor: DR. ZOE KOURTZI

This dissertation is submitted for the degree of MASTER OF PHILOSOPHY IN BIOLOGICAL SCIENCE (PSYCHOLOGY) BY RESEARCH

Lucy Cavendish College

August 14, 2016

Preface

Data collection and interpretation of results was done in collaboration with colleagues in the lab: Ignacio Perez-Pozuelo and Dr. Rui Wang. I confirm that this thesis is my own work except as declared here and that I have documented all sources and materials used. This work was done wholly in candidature for a research degree at this University. Appendix B contains the supplementary methods of a previous study that was submitted for publication in Wang et al. (in review); the original submission was lightly edited to reflect the current study. No other part of this thesis has been previously presented to another examination board or published. This work does not exceed the prescribed word limit of the Degree Committee.

Acknowledgements

Many thanks to my supervisor, Dr. Zoe Kourtzi, for guiding me in this project. Thank you to my examiners, Dr. Lucy Cheke and Dr. Caroline Di Bernardi Luft, for very useful discussion and comments. I am also grateful for the mentorship of Dr. Rui Wang and for helpful discussions with my labmates, especially Joseph Giorgio and Vasilis Karlaftis. Finally, I have had excellent collaborators and thank Ignacio Perez-Pozuelo and Andre Zylstra for their many hours (of work and putting up with me).

I remain ever indebted to the people who helped me get this far: my family, teachers, mentors, and friends.

Lastly, thank you to all of my participants and especially my labmates: Lizzie, Joey, Rui, Nuno, Lukas, Andre, Saana, Vasilis, Poly, Reuben, and Nacho. You've made it a brilliant year.

This work was supported by grants to Dr. Kourtzi from the Biotechnology and Biological Sciences Research Council [H012508], the Leverhulme Trust [RF-2011-378] and the [European Community's] Seventh Framework Programme [FP7/2007-2013] under agreement PITN-GA-2011-290011.

Abstract

We are constantly barraged with information, and yet to navigate the world successfully we must be capable of isolating relevant units. The process of learning statistical relationships between items or events is called statistical learning, and it is a ubiquitous and powerful learning mechanism spanning many modalities and domains. With such a powerful learning mechanism at their disposal, humans must use strategies to direct their abilities in effective manner. We sought to understand how people use two strategies, matching and maximization (loosely analogous to exploration and exploitation respectively), to learn sequences that mimic the complexity of the real-world visual environment. Subjects observed sequences of symbols generated by a first-order Markov process and were asked to predict the upcoming stimulus. Computational modelling was used to extract the learning profiles of participants and the strategies that they followed in learning patterns and making predictions. Using these indices, we tested subjects in seven different conditions, hypothesizing that subjects would change their strategies based on different situational demands. We first tested whether subjects would engage different strategies when directed feedback was provided. We predicted that subjects would use more maximization strategies when task structure encouraged maximization, and this was what was observed. We then tested whether subjects would change strategies when stimuli were presented in an irregular manner. We predicted poorer performance and poorly-performing strategies, but we instead observed more maximizing behaviour. Finally, we tested whether subjects would adopt new strategies when the ease of specific strategies was manipulated. We predicted that more subjects would use a matching strategy in the group with the comparatively easier matching strategy, but in actuality subjects continued to use the same strategies across both groups. This experiment explored and supported the hypothesis that subjects change their strategies with respect to task structure even in a complex paradigm imitating statistical learning in the real world.

Contents

List of Figures

Chapter 1

Introduction

Humans are barraged with complex sensory information in their daily lives. How can they extract relevant units from the overload of stimulation? One powerful mechanism is *statistical learning*, the process of learning the statistical relationships between items and events to understand the structure of the world. For example, consider the case of infants trying to learn language. People deliver a continuous auditory stream, and babies must determine the structure in this sequence to isolate "words". How do babies do this? If an adult says, "the baby laughs" and "the baby eats" for example, the two syllables *"ba"* and *"by"* co-occur more frequently than any other two syllables in these phrases. Babies are sensitive to the transitional probabilities between words and this is one mechanism by which they learn. In fact, seminal studies by Saffran, Aslin, and Newport (1996) showing that infants could learn to segment words by extracting the transitional probabilities between nonsense syllables are the foundation of modern statistical learning.

Statistical learning is a powerful means of extracting information about our environments. The language-learning study by Saffran, Aslin, and Newport (1996) is a specific example of statistical learning and illustrates the probabilistic nature of searching for relationships between neighbouring items (in the case syllables). Since the original language learning studies, the field has expanded in depth and breadth. It spans to encompass a study by Brady and Oliva (2008) showing that after adults observed a stream of 12 scene images organized into triplets by scene category, when they viewed names of those scene categories later (e.g. "Forest", "Mountain") they identified triplets that followed the order of the original images as "more familiar" than triplets with random orderings of the words. Statistical learning has been shown to take place in children and adults (Saffran, E. Johnson, et al., 1999), within a few minutes (Saffran, Aslin, and Newport, 1996; Aslin, Saffran, and Newport, 1998), automatically (Fiser and Aslin, 2001; Brady and Oliva, 2008), without conscious awareness (Kim et al., 2009), across modalities (Conway and Christiansen, 2005), and can transfer across space and time (Turk-Browne and Scholl, 2009). It is a ubiquitous means of segmenting and categorizing regularities from continuous streams of information so that predictions can be accurately made about the world.

Statistical learning is the foundational mechanism, but how do people direct this capacity for probabilistic associations in daily life: what are the strategies that people use in learning complex statistical relationships? The strategies that people use depend on their desired outcomes: people will learn information so that they able to predict future events and ensure the best possible result. Two common strategies that people use are exploration and exploitation. Suppose one is trying to find a best restaurant to eat dinner. One strategy is to return to the family favourite: this restaurant usually produces good food, so one is likely to enjoy a hearty meal. This is an exploitation strategy: capitalizing on something with a known outcome. One could also try a new restaurant, in the hope that it will have even better food. This is an exploration strategy: sampling to determine outcomes, so that better choices can be made in the future. The strategies that people use to navigate a probabilistic world determine their outcomes— in this case, the quality of their meal, but these same strategies apply to higher-stakes decisions like buying a house. Understanding why people use one strategy or another— or influencing which strategies people use in certain situations— is thus an important goal. Maximizing (analogous to exploitation) and matching (analogous to exploration) strategies have been studied scientifically, but often not in the context of statistical learning. Since statistical learning is the mechanism by which most of us learn most things in life, it is key to understand how people direct their learning in choosing strategies to predict and make decisions.

The following experiment aims to address the question of how people use strategies in statistical learning. It would seem useless for humans to not use different strategies in different situations, so we hypothesized that by placing people in different learning environments, they would use different strategies to acquire and predict information. In the first set of experimental groups, we asked whether peoples' strategies would change if we changed their desired outcomes: if we gave them feedback indicating which component of learning was deemed relevant. We expected that subjects would grow to use more maximizing strategies as feedback emphasizing maximization was increased. This effect was indeed observed. In the second set of experimental groups, we asked whether peoples' strategies would change if we altered the temporal dynamics of how information was delivered to them: if we upset the regular temporal spacing between items and instead presented stimuli in an uneven jittered manner. Given the fundamental nature of timing relationships in determining if two events are associated, we expected that subjects' strategies would change with this manipulation— that they would perform poorly in situations with less regularity and adopt strategies with poorer outcomes. In fact, people adopted strategies with *better* outcomes. Finally, in the third set of experimental groups we asked whether peoples' strategies would change if one strategy were made more or less difficult compared to the other. We predicted that subjects would more often choose easier-to-execute strategies independent of the effectiveness of the strategy. We observed that subjects instead were not very much affected by such a manipulation.

In this experiment, we investigated the question of how subjects would use maximizing and matching strategies to direct their statistical learning. We hypothesized that depending on the situational constraints, subjects would differ in their relative uses of these strategies. We implemented feedback manipulations, temporal jitter manipulations, and manipulations changing the structure of the underlying sequence which facilitated the relative use of one strategy over the other. Though matching and maximization strategies have been previously investigated, we sought to determine their use in a complex statistical learning paradigm mimicking the complexity of the decisions people make in everyday life.

1.1 Paradigms

There are many paradigms investigating statistical learning, and some are more useful than others in investigating the question of strategy use and the impact of task structure. Some of these statistical learning paradigms include artificial language learning, visual statistical learning, artificial grammar learning, sequence learning, audio-visual learning, and contextual cuing. All of these paradigms propose to investigate the same abilities, but they differ in their task complexity (and therefore ecological validity), whether individual differences in subject behaviours / strategies can be measured, the precision of response measures, and whether previous generalisation and transfer results have been observed. The latter concept of generalisation / transfer / abstraction goes by many names, but refers to the idea that the rules that people extract about the world should apply to more than specific instances: for example, when babies are learning the word "shirt", they should learn the more general use of the term rather than thinking it means "red shirt" given the two red shirts they have seen. People should then be able to demonstrate their acquisition of such rules by applying them to new domains and modalities. Abstraction often refers to the ability to generalise over surface features like colour or other forms of noise; transfer is the application of learned rules. (See Aslin and Newport (2014) for additional discussion.)

1.1.1 Artificial Language Learning

The original statistical learning studies by Saffran, Aslin, and Newport (1996) involved infants listening to undifferentiated streams of sounds. These streams consisted of triplets of syllables that followed each other continuously: e.g. ***bi**-da-ku-**pa**-do-ti-**go**-la-bu-**bi**-da-ku* where the beginning of each triplet is bolded. The only method to determine where a "word" (triplet) began was to observe the transitional probabilities between sounds: each syllable within a word would follow the preceding syllable with 1.0 probability, but the beginning syllables of words would follow other syllables with probability .33. In the test phase, infants heard two different types of triplets: either syllable triplets in the same order as they had heard in the training phase ("words") or triplets in unfamiliar orders ("non-words"). After only two minutes of training, at test infants spent more time listening to non-words than words, a novelty effect showing that they had attended to the transitional probabilities of the syllables and were able to successfully distinguish between words and non-words. This demonstration of statistical learning was confirmed and expanded upon in follow-up studies (Saffran, Newport, and Aslin, 1996; Saffran, Newport, Aslin, et al., 1997; Saffran, E. Johnson, et al., 1999; Saffran, 2001; Saffran, 2002; Maye, Werker, and Gerken, 2002; Saffran and Thiessen, 2003). The application to ecological language acquisition has been recognized, and the idea that children execute first-order statistical learning (applying the assumption that words, like syllables, are predictive of other words) is computationally supported (Goldwater, Griffiths, and M. Johnson, 2009).

The traditional artificial language learning paradigm has produced powerful results, though it is limited by the use of simple patterns like triplets. Subjects are also limited in the behaviours they can exhibit as looking time is a constrained response measure. This paradigm can produce strong transfer results, however, elucidating the underlying rules that subjects learn (e.g. Marcus, K. Fernandes, and S. Johnson (2007)).

1.1.2 Visual Statistical Learning

Fiser and Aslin (2002a) developed a visual version of the Saffran, Aslin, and Newport (1996) artificial language learning paradigm in which triplets of shapes appeared one by one from behind an occluder (e.g. **A**-B-C-**G**-H-I-**D**-E-F-**A**-B-C. . . , where each letter represents a shape). These triplets were identifiable only by their transitional probabilities, and the response measure was a familiarity test: subjects were required to choose whether the triplets they had observed were more familiar than "foils"— triplets composed of syllables related by a joint probability of zero (e.g. G-A-D). Subjects reliably passed this familiarity test, demonstrating visual statistical learning; these findings have been confirmed in further studies (Fiser and Aslin, 2001; Fiser and Aslin, 2002b; Fiser and Aslin, 2005; Kirkham, Slemmer, and S. Johnson, 2002; Turk-Browne, Jungé, and Scholl, 2005; Turk-Browne, Isola, et al., 2008; Baldwin et al., 2008). Familiarity tests are often used in visual statistical learning studies (Fiser and Aslin, 2001; Fiser and Aslin, 2002a; Turk-Browne, Jungé, and Scholl, 2005), but implicit reaction time measures are also used, under the assumption that if statistical learning took place then subjects will be primed to respond faster to the second and third syllables in a trained triplet (Hunt and Aslin (2001), Turk-Browne, Jungé, and Scholl (2005), and Kim et al. (2009)). Contextual cuing paradigms also use response time (Chun and Jiang, 1998; Olson and Chun, 2001; Brady and Chun, 2007). Response time measures can be more powerful than familiarity measures because they capture whether a subject is *predicting* an upcoming stimulus rather than just classifying it. fMRI response measures have also been applied (Schapiro, Kustner, and Turk-Browne, 2012; Karuza et al., 2013; Schapiro, Gregory, et al., 2014; Turk-Browne, Scholl, et al., 2010).

Typical visual statistical learning studies are also limited by simplistic triplet patterns as well as the individual differences subjects can exhibit within the constraints of familiarity tests and implicit reaction times as response measures. However, these studies have also produced powerful generalisation results across colour, shape, time and space among others (Turk-Browne, Isola, et al., 2008; Turk-Browne and Scholl, 2009).

1.1.3 Artificial Grammar Learning

In an artificial grammar learning paradigm, subjects typically view strings of items, often letters, that have been generated through a finite-state grammar. All of the letters from a string are presented simultaneously, and strings are generally around 3-8 letters long; an example of a grammar and associated strings is shown in Figure 1.1. At test, subjects are exposed to new strings either generated by the grammar ("grammatical") or strings that violate at least one of the rules of the grammar ("ungrammatical"). Subjects who correctly classify grammatical and ungrammatical strings have exhibited statistical learning (Reber, 1967).

The stimuli from artificial grammar learning paradigms are complex and more similar to the type of stimuli generated in our daily environments. These studies do have the limitation that individual differences are difficult to capture due to the one-measure "grammaticality" response. Moreover,

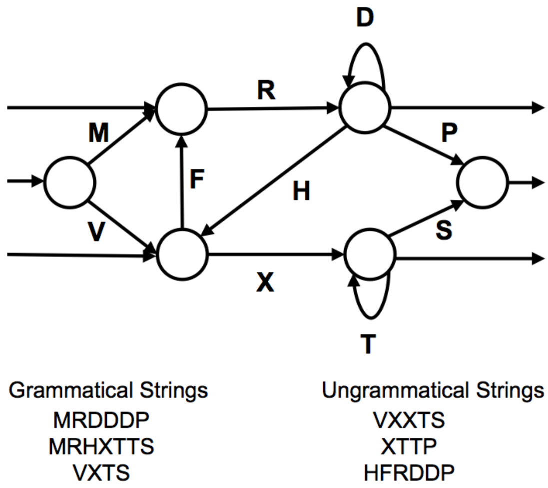


FIGURE 1.1: A finite grammar. Letters are appended to a string.

because the task is complex, it is difficult to determine which aspect of the training stimuli has been learned. To investigate this, the field has distinguished between "non-transfer" learning, which is described in the preceding paragraph, and "transfer" learning, which is the ability to classify strings as grammatical when the strings have been generated using the same grammar but are executed in a new vocabulary (new letters) (Lotz and Kinder, 2006). Artificial grammar learning has thus made progress in trying to isolate what is learned and what can be generalised using a complex stimulus set.

1.1.4 Sequence Learning

Sequence learning predates the modern conception of statistical learning and is distinct in that it uses only deterministic, non-stochastic sequences (see Schwarb and Schumacher (2012) for a review). Sequence learning was developed to study implicit spatial sequence learning and uses the serial reaction time task, developed by Nissen and Bullemer (1987). In Nissen and Bullemer (1987), an asterisk appeared in a position on a screen and subjects pressed a button in the corresponding position. The asterisks appeared in a 10-position sequence (e.g. 4-2-3-1-3-2-4-3-2-1, where each number represents one of the four positions) for the test group, and at random locations for the control group. Subjects in the test group responded more quickly and accurately than those in the control group, demonstrating sequence learning. Sequences such as that used in Nissen and Bullemer (1987) are called hybrid sequences because different positions can follow from any specific position: e.g. "3", "4", and "1" all appear after "2" at specific points.

This sort of sequence is distinguishable from "unique" sequences in which there is a single path (e.g. 4-2-3-1) (Cohen, Ivry, and Keele, 1990).

Sequence learning studies use deterministic sequences, which are not typical of statistical contingencies in the environment. However, one of the great advantages of sequence learning is that individual learning can be accessed online in a continuous way, rather than probed at the end of training during a test phase. Questions can be asked about which specific items are being learned and when they are being learned using the precise, trial-by-trial reaction time measure, and this response moreover requires subjects to predict the upcoming stimulus.

1.1.5 This Study's Paradigm

To answer the question of what strategies subjects use in statistical learning, we sought to develop a paradigm that combined the most relevant elements of the various statistical learning tasks available. Previous studies have combined the serial reaction time task with artificial grammar learning (Cleeremans and McClelland, 1991), language-learning tasks (Hunt and Aslin, 2001; Misyak, Christiansen, and Tomblin, 2010), and tasks with triplet structure (Hunt and Aslin, 2001; Howard et al., 2008). Gómez (1997) used an artificial grammar learning paradigm but instead of having letters in strings displayed simultaneously on a screen, letters were sequentially presented one at a time. Untrained letters were also presented sequentially in the transfer condition of this study. From another direction, Baker et al. (2014) developed a non-serial-reaction-time sequence learning task that incorporated probabilistic elements from visual learning tasks— the sequences were probabilistic but alternated so that the subject faced probabilistic choices. Many studies involved multiple sessions of training taking place over several days (Cleeremans and McClelland, 1991; Hunt and Aslin, 2001; Howard et al., 2008; Baker et al., 2014) with sleep consolidation (Durrant, Taylor, et al., 2011; Durrant, Cairney, and Lewis, 2013), rather than the rapid, minute-long learning that characterizes most statistical learning studies. In this study, subjects trained for five sessions over five days so that their progress could be tracked over time and individual differences could emerge.

The paradigm developed for this study (very similar to Wang et al. (in review)) captured the task complexity of artificial grammar learning, the predictive and trial-by-trial response precision of sequence learning tasks, and the control of transfer conditions exemplified in visual statistical learning. Subjects predicted upcoming symbols while viewing sequences of symbols generated by probabilistic first-order Markov processes (meaning that each symbol depended probabilistically on the immediately preceding symbol). Computational models were developed (originally in Wang et al.

(in review)) to extract several different learning measures and strategy indices, so that the type of learning and decision strategies employed could be tracked for individual subjects on every trial.

Three transfer tests were also developed: an untrained probability transfer test, an untrained symbol transfer test, and an untrained speed transfer test to probe what information subjects extracted and could apply to new sequences. The probability transfer test reversed the relationships governing stimuli order, and asked whether training on the sequence-learning task would facilitate acquisition of new sequences where previously low-probability information was emphasized. The symbols transfer test asked whether subjects had learned the underlying grammar at a level that would enable them to map their learning onto a new set of symbols. The speed test asked if training on the sequence-learning task would enable subjects to learn novel sequences faster than they had been capable of previously. This paradigm lent itself well to the transfer studies and also to interventions, such as feedback manipulations.

The experiment was designed to address the question of what strategies people employ during statistical learning. The study is unique among statistical learning paradigms in that it is specifically adapted to this purpose. The probabilistic and long-term structure of the study mimicked real-life statistical learning and encouraged participants to adopt differentiable and complex strategies. The precise response measures and online-trial-by-trial accumulation of responses allowed the investigation of strategy development over time. The computational models were designed to extract strategy indices for individual participants which allowed comparison between experimental groups. This paradigm incorporated elements of various existing statistical learning studies to address the question of what strategies people use to make decisions in probabilistic environments.

1.2 Strategies

We sought to investigate the strategies that subjects use in statistical learning. What is known about these strategies, specifically maximization and matching? Recall that in the context of choosing a restaurant, exploitation (analogous to maximization) would be choosing a restaurant you already know is good (taking advantage of an option with known rewards), while exploration (analogous to matching) would be taking the risk of trying somewhere new (venturing into the arena of unknown rewards in the hope of optimizing later). Maximizing behaviour is often the optimal behaviour in probabilistic lab settings, where stochasticity is carefully controlled: subjects should find the choice with the highest probability and always pursue that choice. Matching behaviour, on the other hand, is responding in such a way so that subjects' choice probabilities match the outcome probabilities (Koehler and James, 2014). For example, in the paradigm used in this study, "C" follows "B" 80% of the time, but 20% of the time the loop will move backwards and "A" will follow B instead. Subjects who exhibit perfect maximization behaviour will choose "C" 100% of the time. Subjects who exhibit perfect matching behaviour will choose "C" 80% of the time, and "A" 20% of the time, matching the outcome probabilities that they have observed.

Maximizing and matching are strategies for both responding and learning information. If a subject is a matcher they are guaranteed to have learned all of the statistical contingencies between events— they must have learned that "A" follows "B" 20% of the time and that "C" follows "B" 80% of the time. Maximizers, however, can get away with only learning that "C" often follows "B"; they do not necessarily need to know anything about "A". Investigating these strategies thus has important implications surrounding not only how people make decisions but also how people learn. It makes sense to know how to manipulate these strategies so that the appropriate strategies are engaged in the many cases of uncertainty that we face in our daily lives.

If maximization is the easier and the rational response, what is the role of a matching strategy? Ever since this question started being investigated in the 1950s (see Vulkan (2000) for a review), researchers have struggled with why a matching strategy exists at all. In a binary prediction task (explained below), which is the standard task for investigating these strategies, matching violates rational choice. Consider the task of watching an experimenter pull marbles out of a bag, one at a time, and puts them back as soon as she has shown them to you. You have learned that 70% of these marbles are green, and 30% are red. You are to guess the colour of the next marble and are paid a bonus every time you do so correctly. The rational choice would be to choose green 100% of the time, a maximizing strategy, which would ensure you were correct 1.0*.70 + 0.0*.30 = 70% of the time. If you used a matching strategy, you would only be correct .70*.70 + .30*.30 = 58% of the time. The fact that matching behaviour emerges at all— sometimes that it even seems to be the default strategy on these tasks— has puzzled researchers for decades.

Currently the debate has centred around trying to understand whether matching is a fault in our intuitive reasoning systems— matching as a "dumb" strategy and maximizing as a "smart" one— or if matching is an adaptive, sophisticated strategy in specific situations of uncertainty— matching as a "smart" strategy and maximizing as the dumb one (see Koehler and James (2014) for a comprehensive review). From the "matching is dumb" perspective, the strongest argument is that people are rationally guaranteed to earn a higher payout using a maximizing strategy. It is possible that the maximizing strategy simply does not occur to subjects: when Koehler and James (2010) separated their subjects into two groups, one that was explicitly told about the maximizing and matching strategies and asked which would earn them more money, and one group that was not, more subjects in the group given the "hint" maximized. Moreover, West and Stanovich (2003) have shown that subjects who maximize tend to have higher self-reported SAT reasoning scores, a standardised test used to assess cognitive abilities. However, on the "maximizing is dumb" side of the debate, children tend to maximize (Derks and Paclisanu, 1967), subjects under cognitive load tend to maximize (Wolford et al., 2004), and subjects deprived of glucose tend to maximize (McMahon and Scheel, 2010). More relevant to the complex paradigm used in this study, matching behaviour demonstrates that difficult-to-learn low-probability information was acquired, while this is not necessarily the case when subjects demonstrate maximizing behaviour. Of course, subjects could be acquiring all of the low-probability information in their heads and still behaviourally maximizing, a point discussed in Koehler and James (2014) under the name of "pattern matching" and why the exploitation vs. exploration analogy discussed earlier is useful but incorrect to some degree.

Maximization and matching are commonly investigated in simpler learning paradigms— the marble example is a classic case in that the prediction for the next marble is probabilistic (you have a 70% of guessing the next marble) but does not depend on events in the past. In other words, your guess for the next marble does not at all depend on which marble was pulled out previously. In the real world, this kind of sampling independence is not often encountered: what happened yesterday determines where you are today; wherever you are now is not independent of where you were 30 seconds ago. We sought to investigate how subjects used maximization and matching strategies in a context in which previous events mattered. In this experiment, the immediately preceding item probabilistically influenced which item would appear next, a pattern more realistic to the events encountered in daily life.

1.3 This Study

Statistical learning is a powerful mechanism that describes a variety of cognitive functions, spans many paradigms, and uses various response measures. All of statistical learning falls under one title in the expectation that it describes one set of learning mechanisms or processes (Kirkham, Slemmer, and S. Johnson (2002) and Perruchet and Pacton (2006), cf: Bays, Turk-Browne, and Seitz (2016)). As Frost et al. (2015) describe: "Our approach construes statistical learning as involving a set of domain- general neurobiological mechanisms for learning, representation, and processing that detect and encode a range of distributional properties within different modalities or types of input." As a primary means of understanding and interpreting information in our daily lives, it is important to investigate how people employ these learning abilities. Specifically, we asked what strategies subjects use in harnessing their ability to observe and predict probabilistic behaviour.

We hypothesized that under different situational conditions, people would change their use of maximization and matching strategies. When subjects were faced with feedback that changed their desired outcomes, it was predicted that subjects would maximize more when maximizing feedback was employed. When subjects were faced with the task of predicting items that were presented in unusual timing grouping (temporal jitter added to the stimulus presentation), it was predicted that they would perform poorly in the learning task and adopt less optimal— fewer maximizing— strategies. Finally, when subjects encountered sequence structures that made matching strategies easier or harder to use, it was predicted that they would match more when the matching strategy was made easier. This set of experiments addressed the question of how people use strategies in a complex statistical learning paradigm that drew upon principles of everyday learning, and additionally employed transfer tests to understand the content of what subjects learned and how their learning generalised.

This set of experimental groups is described in several chapters, all investigating how strategies changed in different situations. Chapter 3 describes the findings from the Feedback groups, Chapter 4 describes the findings from the Temporal Jitter groups, Chapter 5 describes the findings from the Structural Contingencies groups. In Appendix A we ask about individual differences with group-general findings and cognitive tests. Chapter 6 provides a general discussion of the results of all experiments.

Chapter 2

Methods

All means are reported with standard error.

2.1 Observers

115 naïve volunteers participated in 3 sets of experimental comparisons: Feedback, Temporal Jitter, and Structural Contingencies. There were a total of 7 experimental groups within these 3 sets. 18 volunteers participated in Group 1 ("Main", mean age = 23$\pm$.6 years); 18 volunteers participated in Group 2 ("Feedback", mean age = 22$\pm$.6 years); 14 volunteers participated in Group 3 ("Maximization Feedback", mean age = 23$\pm$.7 years); 14 volunteers participated in Group 4 ("No Feedback", mean age = 22$\pm$1.0 years); 18 volunteers participated in Group 5 ("Jitter", mean age = 23$\pm$.6 years); 19 volunteers participated in Group 6 ("Augmented Jitter", mean age = 23$\pm$.6 years), 1 subject excluded due to noncompliance; 14 volunteers participants participated in Group 7 ("Different Contingencies", mean age = 23$\pm$.8 years). All had normal or corrected-to-normal vision, gave written informed consent and were compensated monetarily (7 pounds/hour) for their time. This study was approved by the Psychology Research Ethics Committee of the University of Cambridge.

2.2 Experimental Setup

Testing was conducted in three dimly lit testing rooms on two monitor types: (Setup 1) a gamma-corrected 21-inch ViewSonic P225f colour monitor (1024 pixel x 768 pixel resolution; 0.38mm x 0.38mm per pixel) and (Setup 2) a gamma-corrected 23.6-inch VIEWPixx / 3D colour monitor (1920 pixel x 1080 pixel resolution). Subjects used a chin rest positioned 60cm (Setup 1) and 43cm (Setup 2) from the monitor to equate viewer stimulus size.

TABLE 2.1: RSVP task parameters.

RSVP	Baseline	Rapid	Group-Specific
# Items per Trial	10	9-13	9-13
Presentation Time (ms)	100	100	100
Inter-Stimulus Interval (ms)	100	100	400
Spatial Jitter (° of visual angle)	0	.6	.6
Temporal Jitter (up to %)	0	0	40 ("Jit."), 60 ("Aug. Jit.")
# Times Target could appear	0-1	0-2	0-2

2.3 Cognitive Assessment Tasks

Independent from the sequence learning task, subjects performed several tasks to assess suitability for inclusion in the sequence learning task and general cognitive performance.

2.3.1 Rapid Serial Visual Presentation (RSVP) task

Subjects were tested on three rapid serial visual presentation (RSVP) tasks ("Baseline", "Rapid", "Group-Specific") to evaluate their abilities to attend to temporally-demanding stimuli (Table 2.1). Subjects were shown a first image, and then a stream of other images, and responded how many times the first image appeared in the stream.

Specifically, RSVP structure was based on (Potter et al., 2002) and code was adapted from `http://web.mit.edu/kehinger/www/PTBexamples.html` (link since removed). Subjects were first shown a fixation cross (500ms) followed by an image designated as the target image (500ms). After a 1s delay, subjects were shown a stream of either 10 different images of familiar objects (Baseline RSVP) or 9-13 different images of unfamiliar symbols (Rapid and Group-Specific RSVP). Familiar object images were downloaded from `http://web.mit.edu/kehinger/www/PTBexamples.html` (link since removed); symbols were taken from the Ndjuká (`http://www.omniglot.com/writing/ndjuka.htm`), Old South Arabian syllabaries (`http://www.unicode.org/charts/PDF/U10A60.pdf`), and Qataban (`http://www.alanwood.net/unicode/fonts-middle-eastern.html#oldsoutharabian`, follow links to Qataban download) and were not used in the sequential learning task. Images were drawn randomly from a pool of 20 alternatives and subtended $8.2°$ of visual angle.

Items in the stream were presented for 100ms with a 1) 100ms inter-stimulus interval (Baseline and Rapid RSVP) or 2) 400ms inter-stimulus interval (Group-Specific RSVP). Random spatial jitter up to $\pm 0.6°$ of visual angle was applied to the Rapid and Group-Specific RSVP tasks; jitter was

centred on the fixation point and occurred in vertical and horizontal axes independently. In the Group-Specific RSVP, subjects that went on to complete the Jitter and Augmented Jitter groups saw inter-stimulus interval blocks with mean 400ms but uniformly and randomly distributed across the values
200/300/400/500/600ms (Jitter) or 100/120/.../680/700ms (Augmented Jitter). The target item could appear once in the stream, twice in the stream, or not at all. In the Baseline RSVP, the target object could appear 0-1 times; in the Rapid and Group-Specific RSVPs the target symbol could appear 0-2 times. All other items were not repeated. After the stream of either 10 (Baseline RSVP) or 9-13 (Rapid and Group-Specific RSVP) items, subjects were presented with a 2s window in which they were required to make a 2- or 3-alternative forced choice via keyboard button press indicating how many times the target item had appeared in the stream. This constituted a trial. The next trial commenced after 200ms. Each task was performed once and consisted of 42 trials; participant accuracy was recorded.

Subjects unable to score higher than 80% on the Baseline RSVP task were excluded from participating.

2.3.2 Visual Short-Term Memory (VSTM) task

Participant working memory was assessed with a sequential visual short-term memory (VSTM) task adapted from Luck and Vogel (1997). Coloured dots (diameter $1.7°$ of visual angle) were presented against a grey background for 500ms, followed by an inter-stimulus interval of 1000ms. Subjects were then shown a new screen with the same configuration of dots; dots could be the same colours as previously or changed. A white box appeared around one of the dots (determined randomly) and subjects were required to indicate whether the dot in the box had changed colour via button press. The number of dots displayed was determined via a two-down one-up staircase design to achieve 70% performance. Each block consisted of 10 staircase reversals and the task was performed three times in succession. Working memory threshold for each block was defined as the mean of the last two-third reversals in each staircase. Participant working memory score was recorded as the mean threshold across the three attempts (higher scores indicate better performance).

Subjects unable to average higher than 3.0 on the VSTM task were excluded from participating.

2.3.3 Useful Field of View (UFOV) task

The Useful Field of View (UFOV, Visual Awareness Inc.) task was used to assess selective attention capacity (e.g. described in Edwards et al. (2006)).

Each trial began with a white fixation box presented for 1 second against a black background. The test stimuli were then presented for a variable time as determined by participant ability. A visual mask was shown for 1s to control for afterimage effects, then the response screen was shown. Responses were made using mouse clicks. There were two components to the response: the first tested processing speed and required participants to correctly identify a silhouette (1.9 x $1.4°$ angle) of a car or truck which had been presented centrally inside a white bounding square ($2.9°$ of visual angle). The second component required the participant to identify the location of a simultaneously presented silhouette of a car (1.9 x $1.4°$ of visual angle) at one of 8 radial locations (fixed $10.4°$ of visual angle from central stimulus). Selective attention was assessed by requiring the participant to ignore 47 triangles of the same size and luminance as the radial target. A double staircase was used to determine the display duration at which each participant correctly performed the selective attention task 75% of the time (lower scores indicate better performance).

Subjects unable to average lower than 300ms on the UFOV tasks were excluded from participating.

Additional cognitive tasks described in Appendix A.

2.4 Sequence Learning Task

2.4.1 Stimuli

Three sets of symbols (sets A, B, and C) were generated from the Ndjuká, Old South Arabian and Qataban syllabaries (Figure 2.1); symbols were unfamiliar to participants and highly discriminable. Symbol sets were randomly allocated to participants in the "training" condition (this set was also used in the test and probability transfer test), symbol transfer test, and speed transfer test.

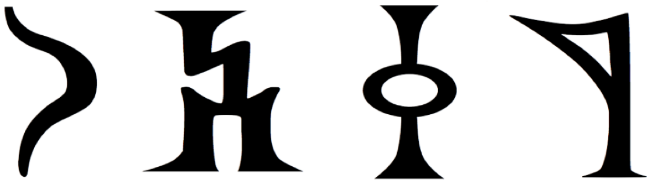
Symbol mappings (which symbol images were assigned to "A","B","C", and "D" in sequences) were randomised across participants. Symbols subtended $8.2°$ of visual angle and were presented centrally in black on a grey background. Random spatial jitter of up to $\pm 0.6°$ visual angle was applied to the presented symbols and occurred independently in both the vertical and horizontal dimensions. Sequences were presented using the Psychophysics toolbox 3 for Matlab (http://psychtoolbox.org) (Brainard, 1997; Pelli, 1997).

2.4.2 Sequence Design

Markov models can be used to generate sequences where individual items depend only on the previous item (first-order Markov models, memory

(A) Symbol Set A

(B) Symbol Set B



(C) Symbol Set C

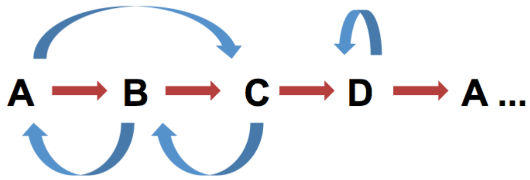
FIGURE 2.1: Symbol sets. Symbol sets A, B and C were taken from the Ndjuká, Old South Arabian and Qataban syllabaries respectively.

length of 1) or do not depend on any previous item (zero-order Markov models, memory length of 0).

In this study a first-order Markov model was used to produce probabilistic sequences of symbols, so any symbol that was presented was dependent on the previously-presented symbol. This previously-presented symbol was called the "context". Within each context, two possible symbols could follow in the sequence, one with high probability (80%) and one with low probability (20%) (Figure 5.1). Each symbol had the same average likelihood of appearing throughout the sequence (the marginal probabilities of all four symbols was fixed at 25%).

Random trials were generated using a zero-order Markov model. In random trials, each symbol had the same average likelihood of appearing throughout the sequence (the marginal probabilities for all four symbols was fixed at 25%), and each symbol had the same likelihood of appearing when context was taken into account (the conditional probabilities for all four symbols was fixed at 25%).

One long sequence was generated each for 1) training, 2) test, 3) random, 4) the probability transfer test, 5) the symbol transfer test, and 6) the speed transfer test. All sequences were periodically interrupted by a response grid, when subjects had to predict the next symbol to appear (the correct answer was the symbol that had been replaced by the response grid,



Level 1		Target				P (c)
		A	B	C	D	
Context	A		0.8	0.2		**0.25**
	B	0.2		0.8		**0.25**
	C		0.2		0.8	**0.25**
	D	0.8			0.2	**0.25**

FIGURE 2.2: Main Markov model design. Red arrows indicate the most likely symbol to appear given the context (80%); blue arrows indicate the less likely symbol (20%). P(c) refers to marginal probabilities. "A", "B", "C", and "D" are replaced by symbols (randomized by participant). This Markov model was used for all groups except Different Contingencies.

which subjects never observed). These sequences were also segmented into "blocks" after which feedback would be given before the sequence continued.

2.4.3 Experimental Groups

Subjects were assigned to three sets of experimental groups (7 groups in total). First, a "Feedback Groups" set contained the control experiment, "Main" and three feedback conditions ("Feedback", "Maximization Feedback", and "No Feedback"). Second, a "Jitter Groups" set contained the control experiment, "Main", and two jitter conditions ("Jitter" and "Augmented Jitter"). Third, a "Structural Contingencies Groups" set contained the control experiment, "Main" and one condition with sequences generated from an alternative Markov sequence ("Different Contingencies"). All sets except for the "Different Contingencies" consisted of sequences generated from the first-order Markov model described in Figure 5.1(a).

Most experiments contained "block feedback": feedback shown to subjects at the end of each block as "Matching Performance Index", a measure capturing how closely the probability distribution of the participant responses matched the probability distribution of the presented symbols. Block feedback was independent of whether subjects chose the *correct* next symbol in the sequence during the prediction task. Rather, a matching performance index of 100% could be reached if subjects matched the probabilities of each symbol given the context. (For example, symbol "A" might appear after "B" roughly 20% of the time, and symbol "C" might appear after "B" 80% of the time. If subjects chose "A" 20% and "C" 80% of the time after seeing "B", and did this probability matching correctly for all four contexts, they would receive a perfect score regardless of the specific items in the sequence hidden by the response grid.)

In each group, subjects underwent test trials, training trials, and transfer test trials. Each training block consisted of 60 trials on a single symbol set, with manipulations assigned based on the group (see Table 5.1). During test segments, subjects completed a test block (40 trials), a random block (40 trials), and another test block (40 trials) using the trained symbol set. Associated experimental manipulations were applied; however, no feedback was given in any test segments.

Transfer test blocks in the final day were 60 trials and consisted of the following in order: "untrained probabilities", "untrained symbols", and "untrained speed". No feedback (block or trial-by-trial) was given during the transfer trials. In the "probability transfer" test, trained symbols were used but the underlying context probabilities were reversed, such that the high-probability symbol in a given context became the low-probability symbol and vice versa. In the "symbol transfer" test, the symbol set was switched (symbol sets were randomly assigned to subjects, so a novel symbol set was selected) but the underlying model for context probabilities was maintained. In the "speed transfer" test, the symbol set was switched to a novel set again. Additionally, symbols were presented as usual for 100ms, but the inter-stimulus interval was shortened from 400ms (in the training condition) to 100ms. No temporal jitter was applied during the speed transfer trials.

2.4.4 Procedure

Subjects who passed the screening tasks were called back for the sequence learning task, which took place over five consecutive days. On Day 1, subjects were initially familiarized with the sequence learning task through a brief practice block containing 20 trials of random sequence. Subjects then completed a pre-training test sequence (40 trials of training sequence / 40 trials of random sequence / 40 trials of training sequence), two pre-training speed transfer test blocks (60 trials each), followed by two training blocks (60 trials each). During Days 2–4 subjects completed 7 training blocks per day. On Day 5 subjects completed a post-training test sequence (40 trials of training sequence / 40 trials of random sequence / 40 trials of training sequence), two probability transfer test blocks, two symbol transfer test

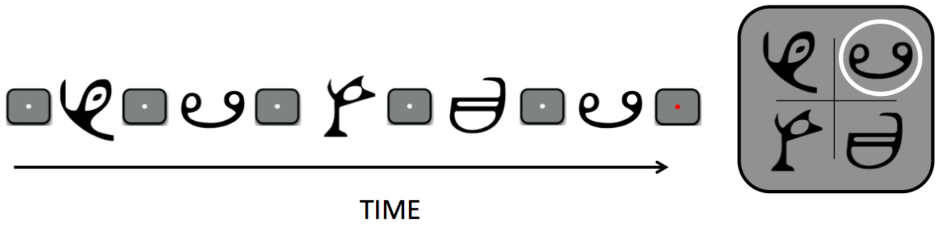


FIGURE 2.3: An example trial. Symbols appeared one by one separated by white dots. At the end of the trial a red dot appeared, after which the grid of the symbols was displayed. Subjects made their decision by keyboard press, and a white circle appeared around their choice. This circle was coloured red or green in the Feedback and Maximization Feedback groups.

blocks, and two speed transfer test blocks. Subjects completed these tests in the fixed order specified. Subjects participated between 45–75 minutes per day.

Each trial consisted of the following. 9-13 symbols drawn from the same symbol set appeared one by one in the centre of the screen in a continuous stream. Each symbol was presented for 100ms and followed by a white fixation dot for an ISI of 400ms (100ms during the speed transfer test). A red dot replaced the white dot at the end of the trial, and was followed by a 2x2 response grid displaying all four symbols in the symbol set. Symbols appeared in randomized locations in the grid. Subjects had two seconds to select which symbol they believed would appeared next with a key press, at which point their selection would be circled (in white in most of the groups). This circle was green (correct) or red (incorrect) in the Feedback and Maximization Feedback groups, based on whichever symbol was hidden beneath the response grid (Feedback) or the most likely choice given the context (Maximization Feedback). After a response had been selected, a fixation dot appeared for 150ms (ITI) before the next trial began. 40 or 60 trials comprised a block. At the end of each block a break of at least 1 minute was enforced (Figure 2.3).

2.5 Data Analysis

2.5.1 Matching Performance Index

Matching performance index is a measure capturing how closely the probability distribution of the participant responses matched the probability distribution of the presented symbols. Matching performance index was calculated for each block. To calculate this index per block, the probability distributions of both 1) the participant responses and 2) the presented symbols were first calculated for each context. Specifically, for each context (e.g.

the previously-presented symbol was "A"), the number of times each symbol was 1) chosen / 2) presented was summed across all of the trials in the block and divided by the number of trials in the block. Once these probability distributions for both 1) the participant responses and 2) the presented symbols were calculated, they were compared by quantifying the amount of overlap between the two distributions: for each symbol the minimum of two probability distributions was taken, and then the values calculated for all four symbols were summed. This gave the matching performance index per context. Finally, the final matching performance index was calculated by averaging the performance indices for each context.

2.5.2 Weak Learners

Subjects with mean normalized post-training test matching performance index scores (mean for both test blocks of post-training test performance − simulated performance assuming random guessing) <.05 were classified as weak learners and were excluded from further analysis.

2.5.3 Modelling Approach

Two sources of knowledge were required to correctly perform the sequence learning task. First, at the most basic level subjects had to be aware that they were viewing a sequence generated by a first-order Markov process rather than a zero-order Markov process: subjects must have learned that any symbol they saw depended on the previous symbol. Second, after this was understood subjects must have learned the specific statistical contingencies between individual items: e.g., given the last symbol shown was "A", the next symbol was most likely to be "B". Subjects' knowledge was extracted based on their response patterns.

Specifically, these two sources of knowledge were measured using a 1) context-length model (probing whether subjects understood that the symbols were generated using a first-order Markov model rather than a zero-order Markov model), and a 2) predictive-contingency model (probing what knowledge subjects had of the statistical contingencies) (see Appendix B, textciteWang2016). The predictive-contingency model was then fed into a 3) strategy model to determine whether subjects were using a maximization strategy rather than a matching strategy in their responses. These were Bayesian models that dynamically updated the weights of evidence that 1) subjects were responding using a first-order Markov assumption rather than a zero-order Markov assumption, or 3) subjects were using a maximization strategy rather than a matching strategy in their responses. The resulting models showed the time courses of how individual subjects 1) learned that sequences were first-order rather than zero-order, and 3) adopted different maximization / matching strategies. The context-length model produces two summary statistics: the learning rate and transition time. The strategy model produces one summary statistic: ICD (and ICD end).

Context-Length Model

The context-length model was used to probe whether participants were using a first-order or a zero-order model to generate their responses during training. This response-tracking model was based on a weighted combination of Markov processes (i.e. first- and zero-order). Mixture coefficient weights were assigned to first-order and zero-order Markov processes; initially, the coefficient of level-1 was set to .2 and the coefficient of level-0 was set to .8, to reflect that subjects were likely to initially assume a simpler model. Participants then began training, and the mixture coefficients and the mixture components themselves were updated after each participant response. The model calculates whether participants' responses were more likely to have been driven by a particular mixture component (e.g. a first-order Markov model) and updates the weights for the components accordingly in a Bayesian manner. (Further details of the model are described in Appendix B, Wang et al. (in review).)

Mixture coefficient curves show the context length that individual participants were using to respond during training. Values of 1 indicate that the subject was more likely to be using a first-order Markov model to generate their responses. Values of 0 indicate that the subject was more likely to be using a zero-order Markov model to generate their responses. Intermediate values indicate that subject were using context lengths that were a mixture between Markov level-1 and Markov level-0. These *learning curves* follow smooth sigmoid shapes. The slope of a subject's learning curve describes how quickly a subject made the transition from responding based on a first-order Markov model compared to a zero-order Markov model, and is called the *learning rate*. The time point (in terms of trials) at which the mixture coefficient weights for level-1 and level-0 both equal .5 captures how long it took for subjects to make this transition, and is called the *transition time*.

Strategy Model

Because sequences were probabilistically generated, subjects needed to learn the statistical contingencies between different symbols in order to make correct predictions in the task. Two common strategies were formulated for how subjects would make their responses. The first strategy was probability *maximization*: choosing the most likely symbol that would appear in a given context. For example, if "A" was the previously-presented symbol (context), and "B" was produced 80% of the time and "D" was produced 20% of the time given this context, then a maximizing response would be to choose "B" 100% of the time. The other strategy was probability *matching*: responding to produce the probability distribution for a given context. For example, if "A" was the previously-presented symbol, then subjects would choose "B" 80% of the time and "D" 20% of the time to match the probabilities which they had observed these symbols appeared given the context. The tendency for subjects to choose a maximizing or matching strategy was quantified with a strategy index, described below.

To estimate the strategies that participants were using in their predictions, individual participants' predictive contingency models (see Appendix B, Wang et al. (in review) for details) were compared to two baseline models: (i) probability matching, a model whose probability distributions derive from the Markov models used to generate the presented sequences (*Model-matching*) and (ii) a probability maximization model, a model where only the most likely symbol was permitted given each context (*Model-maximization*). Kullback-Leiber (KL) divergence was used to compare the participant response distribution to the two models.

The difference between the KL divergence from the participant's predictive contingency model to *Model-matching* and the KL divergence from the predictive contingency model to *Model-maximization* was quantified as strategy choice: ΔKL(*Model-maximization*, *Model-matching*). Negative strategy choice values indicate a strategy more similar to maximization, while more positive strategy choice values indicate a strategy more similar to matching. Strategy choice was computed trial-by-trial and resulted in a *strategy curve* for each participant. Two other artificial data sets were generated simulating responses based on exact matching or maximization.

A strategy index (integral curve difference: ICD) was calculated for each participant based on their strategy curve. Specifically, the integral of each participant's strategy curve subtracted from the exact matching curve as defined by *Model-matching* was taken. This value was calculated across all of training (*ICD*) and for the last two blocks of training (*ICD end*). This integral curve difference between individual strategy and exact matching is close to zero when subjects use strategies similar to matching, and more positive when subjects use strategies more similar to maximization.

Chapter 3

Feedback

3.1 Introduction

In this section, we investigated whether people would change their learning strategies in situations with different desired outcomes. We hypothesized that subjects would change their strategies if faced with different task demands. Feedback, an important element in learning new tasks in daily life (Dale and Christiansen, 2004), details what information is important and should be attended to. The effect of feedback has been investigated broadly (see Kluger and DeNisi (1996) for a review) and specifically in regards to maximization and matching.

Feedback generally seems to encourage maximization: in an artificial grammar learning task (Dale and Christiansen, 2004), even when feedback is completely uninformative (Newell and Rakow, 2007), in addition to the effect of a hint (Newell, Koehler, et al., 2013), and alongside the fact that training can encourage maximization (Shanks, Tunney, and McCarthy, 2002). Maximization is the most rational outcome in uncertain probabilistic tasks, and it is intriguing that feedback seems to encourage people to adopt maximization behaviour under so many circumstances in which it might not be expected to have an effect. It seems that explicitly informing subjects that they are being monitored encourages subjects to maximize rewards in a way they do not when performing "unobserved"— perhaps the reminder of doing an experimental task forces subjects to evaluate what task demands they think the experimenter wishes for them to meet. Initially subjects might not be focused on getting the highest score, but rather exploring the space of options, until they are reminded by feedback that the experimenters may want them to maximize their scores. In this experiment, we asked whether the same pattern of maximization would occur when feedback was given in a more complex paradigm with ambiguous task demands, and what the effect would be of feedback that was informative and directed. Would feedback still result in maximization in a complex paradigm? If so, if directed feedback was given to encourage matching behaviour, would the default to maximizing behaviour occur regardless?

The goal of this study was to ask whether feedback would induce people to change the strategies they used to direct their statistical learning. Four feedback experimental groups were compared to do so. The first group, "Main", included feedback at the end of every block (60 trials) based on whether subjects exhibited matching behaviour ("block feedback"). Block feedback was given in terms of matching performance index, and subjects could receive a perfect score if their response probabilities matched the outcome probabilities irrespective of whether their predictions were correct within the sequence. The second group, "Feedback", included block feedback and trial-by-trial accuracy-based feedback: subjects' responses were circled red or green depending on whether they correctly predicted the next symbol in the sequence. The third group, "Maximization Feedback", included block feedback and trial-by-trial maximization feedback: subjects' responses were circled red or green depending on whether they maximized in the given context or not. The fourth group, "No Feedback", was a control group with neither block nor trial-by-trial feedback.

Given the impact of feedback in simpler domains in the literature, it was predicted that subjects would maximize in the conditions where any feedback occurred: Main, Feedback, and Maximization Feedback. (Alternatively, subjects in the Main condition were predicted to do more matching because of the directed block feedback encouraging matching.) The No Feedback condition was expected to have the least amount of maximization because it included neither block nor trial-by-trial feedback. Since directed feedback in the Maximization Feedback condition was explicitly enforcing maximization behaviour, this condition was expected to have the most maximizers.

The Feedback condition included trial-by-trial feedback as well as block feedback and so was expected to encourage more maximizing behaviour than the Main condition based on the idea that more feedback encourages more maximization. Additional support for the idea that the Feedback group would have more maximizers than the Main group derives from the fact that Feedback participants were receiving accuracy-based feedback, which implicitly encourages maximization since a maximization strategy results in the most correct answers. Further evidence for Feedback having more maximizers and Main having more matchers comes from Gao and Corter (2015), who found that when subjects were rewarded if they perfectly predicted sequences rather than individual trials in a binary prediction task, more matching behaviour occurred. Subjects predicted sequences in the Main group, where they were given block feedback, and subjects predicted individual trials in the Feedback group, where they were given trial-by-trial feedback, so it was expected that more matchers would appear in the Main group and more maximizers in the Feedback group. Gao and Corter (2015) hypothesize this strategy effect occurs because subjects trying to perfectly predict longer sequences as opposed to trials adopt a goal of perfect prediction. To achieve perfect prediction, subjects cannot use maximization, which achieves the highest possible score when the predicted items are random but cannot achieve the 100% score achievable when predicted items are deterministic. We sought to ask whether subjects would adopt similar strategies in a task more complicated than binary prediction.

With regards to hypotheses for the specific measures in this paradigm, two different types of information could be extracted in this study: learning profiles and strategy. Due to the slow time course of learning in this paradigm, subjects in all groups were expected to learn at approximately the same rate, since maximization and matching strategies result in roughly the same learning profiles. Thus the learning profile measures— matching performance index averages, learning rate, and transition times— were expected to be similar across groups. Similarly, subjects were not expected to generalise their performance differently across groups, so matching performance index scores across all transfer tests were expected to be similar.

However, it was predicted that the feedback interventions would change subjects' use of strategies even in the complex paradigm, indicating that subjects adjusted how they were searching for information and responding. Strategy indices are measured in ICD (integral curve difference) and ICD end values, where higher values indicate maximizing behaviour. It was expected that Feedback and Maximization Feedback interventions would push subjects towards maximization behaviour (higher ICD and ICD end scores), more so in the Maximizing Feedback case. The least amount of maximization was expected in the No Feedback group, and so ICD and ICD end scores were expected to be lowest in that group. The Main group was expected to have higher ICD and ICD end scores than the No Feedback group for two reasons: first because any feedback would likely encourage maximization, and second because subjects were receiving matching performance index scores: directed feedback encouraging matching behaviour. Matching behaviour produces higher ICD and ICD end scores than not pursuing either the matching or maximization strategy, so encouraging matching behaviour was also expected to improve ICD and ICD scores for the Main group compared to the No Feedback group.

3.2 Methods

Four experimental groups were compared: "Main", "Feedback", "Maximization Feedback" and "No Feedback". In the Feedback and Maximization

TABLE 3.1: Feedback Experimental Groups Summary. Groups are in the following order: Main, Feedback (FB), Maximization Feedback (Max. FB), No Feedback (No FB). Abbreviations: "Acc." refers to accuracy-based feedback, "Max." refers to maximization-based feedback.

Groups:	**Main**	**FB**	**Max. FB**	**No FB**
Temporal Jitter (up to %)	0	0	0	0
Markov Model	main	main	main	main
Trial-By-Trial Feedback	n/a	Acc.	Max.	n/a
Block Feedback	yes	yes	yes	n/a

Feedback groups where trial-by-trial feedback was given, the circle that appeared when subjects made their prediction for the upcoming symbol on each trial was coloured either green or red ("correct" / "incorrect"). In the Feedback group, feedback was determined by the identity ("A", "B", "C", or "D") of the next symbol in the sequence ("accuracy-based feedback"). In the Maximization Feedback group, feedback was determined by the most-likely symbol to appear based on the context probabilities ("maximization-based feedback"). In the No Feedback group, neither trial-by-trial nor block feedback was given (Table 5.1).

Subjects were trained over four days, and on the last day did a testing sequence (test block, random block, test block), the probability transfer test (the high- and low-probability statistical contingencies were switched), the symbol transfer test (symbols were replaced by a new set of symbols), and the speed transfer test (items were presented more quickly).

3.3 Results

3.3.1 Learning Profile

Matching Performance Index

Subjects had similar learning profiles across the different feedback interventions. Matching performance index patterns across training were similar across groups (Figure 3.1). Mean test scores (pre- and post-) were not significantly different across feedback groups (mixed two-way ANOVA, Session x Group, $F(3,53) = 1.45$, $p = .24$). There was not a main effect of Group ($F(3,53) = .71$, $p = .55$). There was a significant main effect of Session, $F(1,53) = 370.06$, $p < .001$ (Figure 3.2), which indicated that subjects improved from pre-training to post-training.

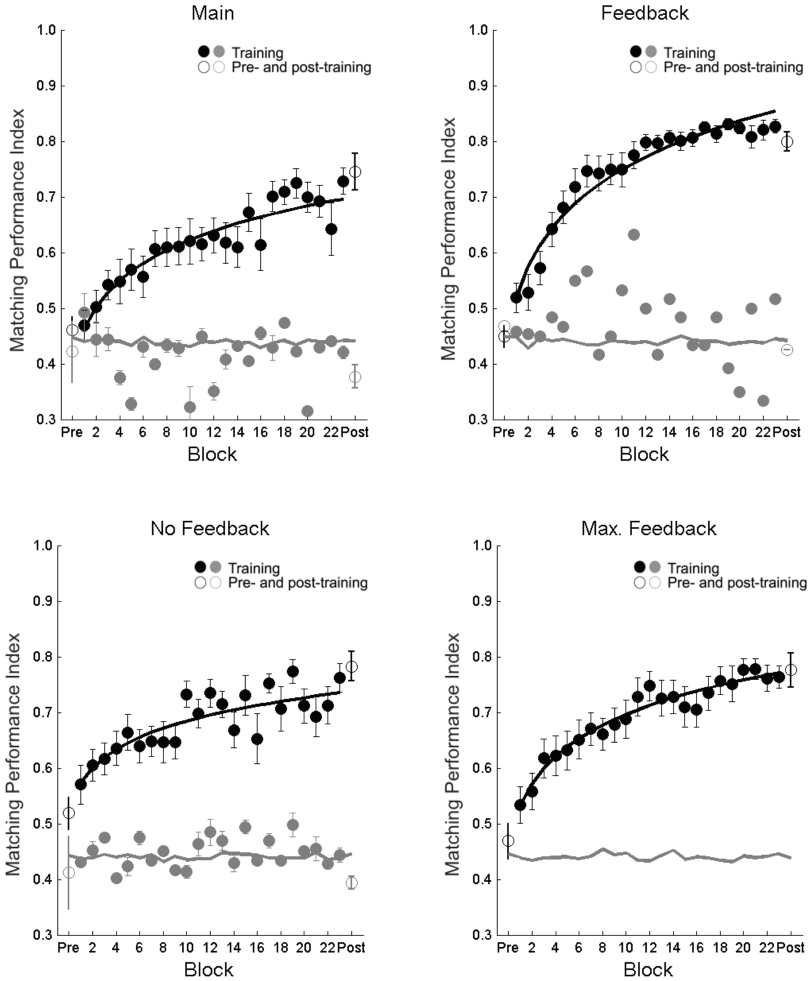


FIGURE 3.1: Behavioural performance. Matching performance index for participants from the Main (n = 15, weak learners: n = 3), Feedback (n = 17, weak learners: n = 1), Maximization Feedback (n = 14, weak learners: = 0), and No Feedback (n = 11, weak learners: n = 3) groups across training (solid circles), the pre-training test (open circles) and the post-training test (open circles). Higher scores on the matching performance index indicates more matching behaviour. Participants completed the task over five days. Data is fitted for participants who improved during training (black circles). Data is also shown for participants that did not improve during training (grey circles). Random guess baseline is indicated by a solid grey line across blocks. Error bars indicate standard error of the mean.

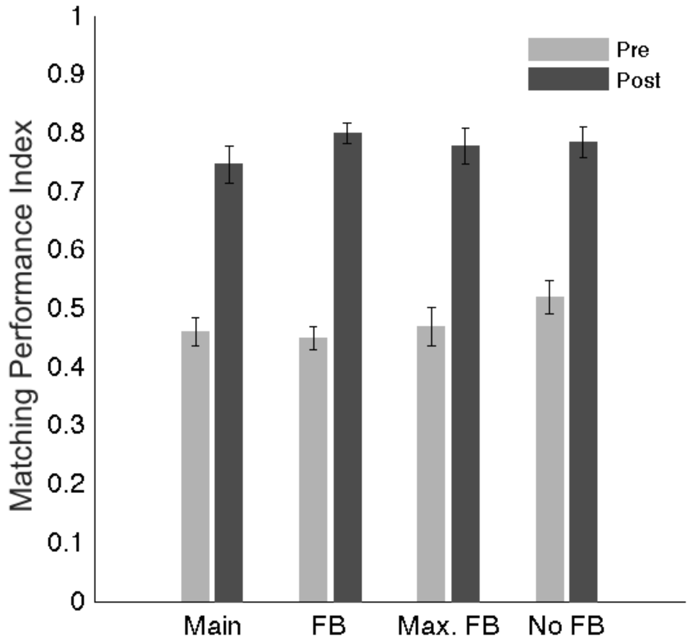


FIGURE 3.2: Behavioural test performance. Matching performance index for participants from the Main (n = 15), Feedback (n = 17), Maximization Feedback (n = 14), and No Feedback (n = 11) groups for pre-training performance and post-training performance. Higher scores on the matching performance index indicates more matching behaviour.

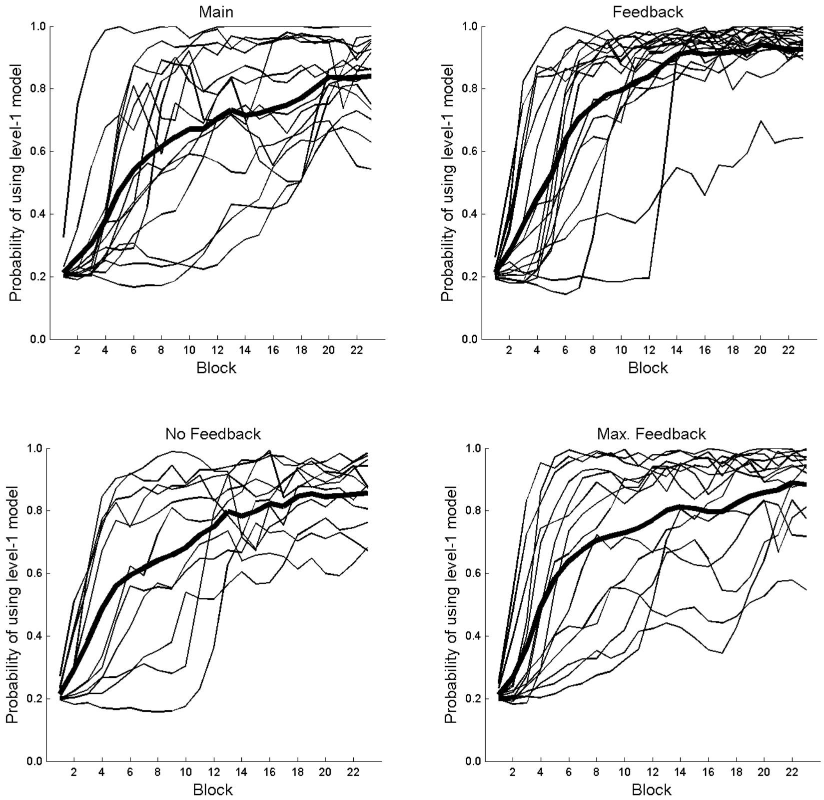


FIGURE 3.3: Learning curves. Mixture coefficient weights for level-1 compared to level-0 model for participants from Main (n = 15), Feedback (n = 17), Maximization Feedback (n = 14), and No Feedback (n = 11) groups. The average curve is shown as a thicker line. Weights closer to 1 indicate that the subject is more likely to be making predictions based on a Markov level-1 model; weights closer to 0 indicate that the subject is more likely to be making predictions based on a Markov level-0 model.

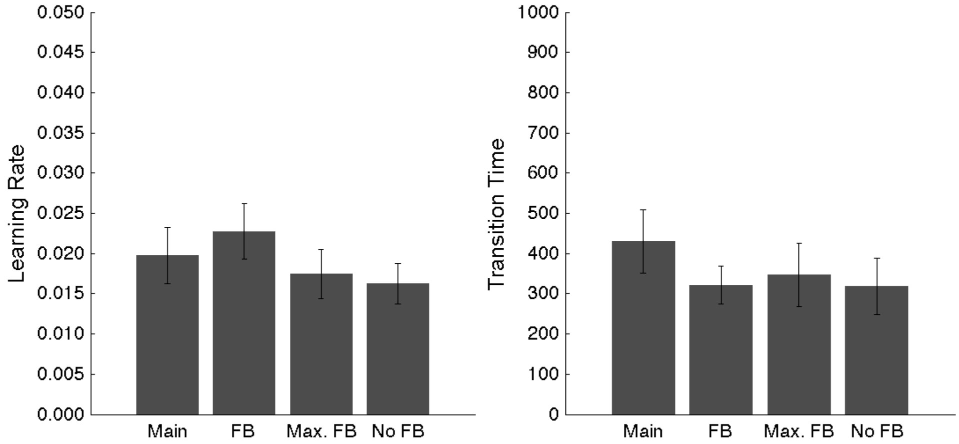


FIGURE 3.4: Mean learning indices for participants from Main (n = 15), Feedback (n = 17), Maximization Feedback (n = 14), and No Feedback (n = 11) groups. Learning rate is the slope of the learning sigmoid curve; transition time is the time point at which y-axis of the learning curves is equal to .5 (the weights for level-1 and level-0 Markov models are both equal to .5). Higher learning rates indicate a faster rate of learning; lower transition time values indicate earlier learning.

Learning Indices

Learning profiles appeared similar across groups (Figure 3.3). Mean learning rates were not significantly different across groups (one-way ANOVA, $F(3,53) = .76$, $p = .52$), so subjects did not learn faster in different groups (Figure 3.4). Mean transition times were not significantly different across groups (one-way ANOVA, $F(3,53) = .60$, $p = .62$), so subjects did not learn earlier in different groups (Figure 3.4).

Strategy

Strategy profiles were affected by feedback manipulations across groups (Figure 3.5). ICD scores were significantly different across groups (one-way ANOVA, mean square between groups = .132, $F(3,53) = 2.80$, $p = .049$). Post-hoc Tukey HSD tests showed no significant differences between individual groups, but narrowly missed significance (Main vs. FB: $p = .068$, FB vs No FB: $p = .097$) (Figure 3.6). ICD end scores were significantly different across groups (one-way ANOVA, mean square between groups = .264, $F(3,53) = 2.80$, $p = .049$). Post-hoc Tukey HSD tests showed no significant differences between individual groups, but narrowly missed significance (Main vs. FB: $p = .082$; FB vs. No FB: $p = .091$.) (Figure 3.6). Thus strategy differed across groups but did not significantly differ between any two specific groups. However, the greatest differences between two groups were between the

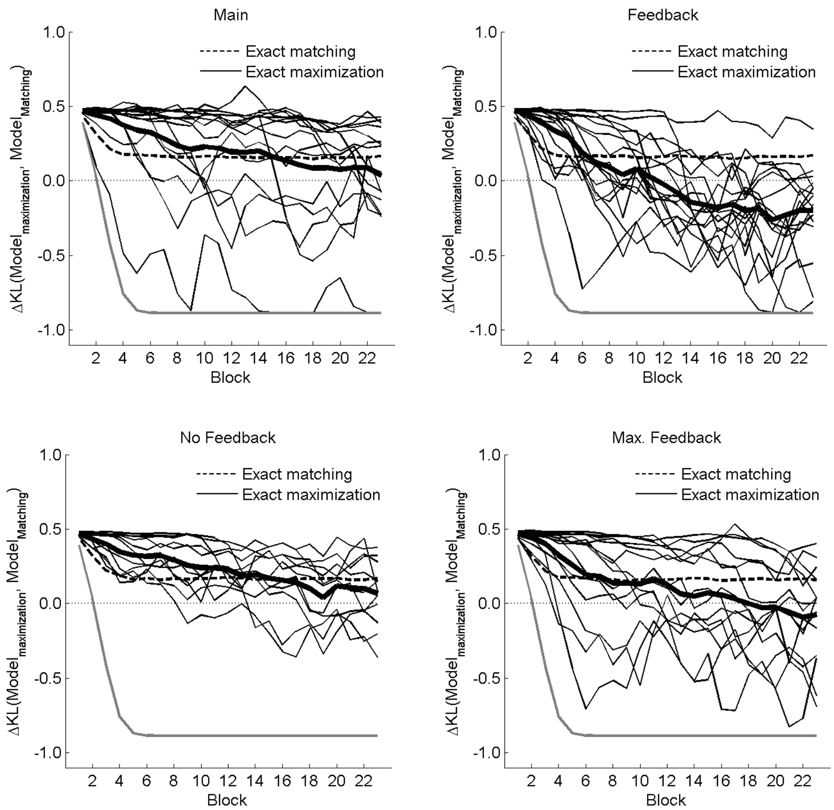


FIGURE 3.5: Strategy choice. ΔKL divergence between model matching and model maximization strategies for participants from the Main (n = 15), Feedback (n = 17), Maximization Feedback (n = 14), and No Feedback (n = 11) groups. KL divergence is a measure of how different the probability distributions that the subject is using to make predictions are from the probability distributions of predictions made using a perfect maximization strategy and a perfect matching strategy. The average curve is shown as a thicker line. Exact matching (dashed line) and maximization models (solid grey line) are plotted. More negative KL scores indicate a strategy closer to maximization.

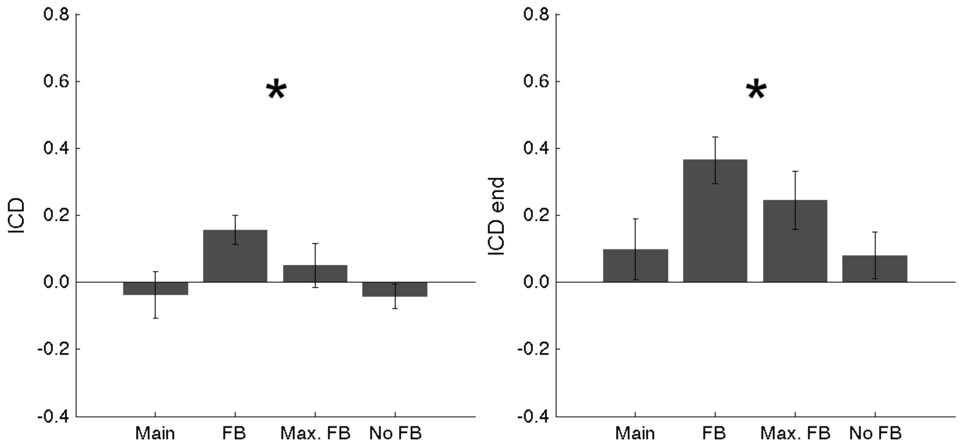


FIGURE 3.6: Mean strategy indices for participants from the Main (n = 15), Feedback (n = 17), Maximization Feedback (n = 14), and No Feedback (n = 11) groups. ICD (integral curve difference) and ICD end measure the signed area between the subjects' strategy curve and predictions made using a perfect matching strategy. Higher ICD and ICD end values indicate a strategy closer to maximization. ICD and ICD end values equal to zero indicate a perfect matching strategy.

Main and No Feedback conditions and the Feedback condition; the most maximization occurred in the Feedback condition.

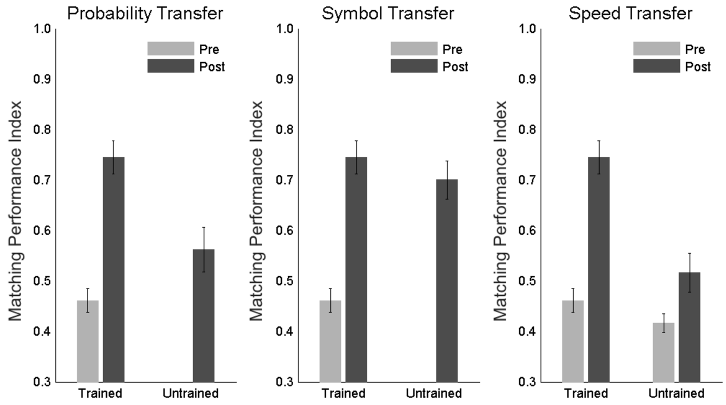
3.3.2 Transfer

Probability Transfer Test

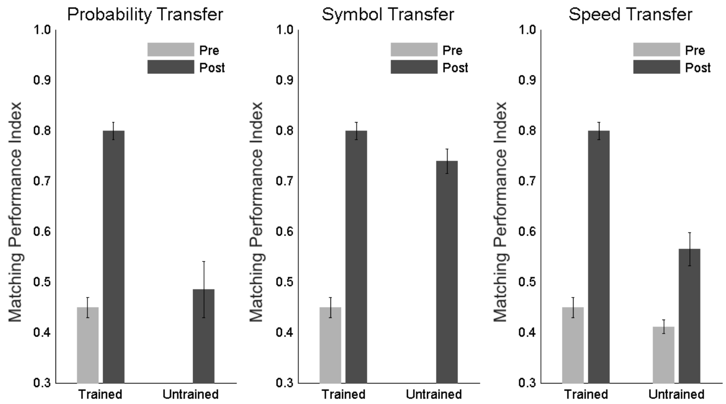
In the probability transfer test the high- and low-probability statistical contingencies were exchanged. Improvement from the pre-training test to the untrained transfer test occurred (mixed two-way ANOVA, significant main effect of Training, $F(1,53) = 7.90$, $p = .01$), showing some generalisation occurred from training to the probability transfer condition. There was not an interaction effect of Training x Group, $F(1,53) = 1.10$, $p = .36$, nor a main effect of Group, $F(3,53)= .85$, $p = .48$, indicating that performance did not significantly differ between groups (Figure 3.7).

Symbol Transfer Test

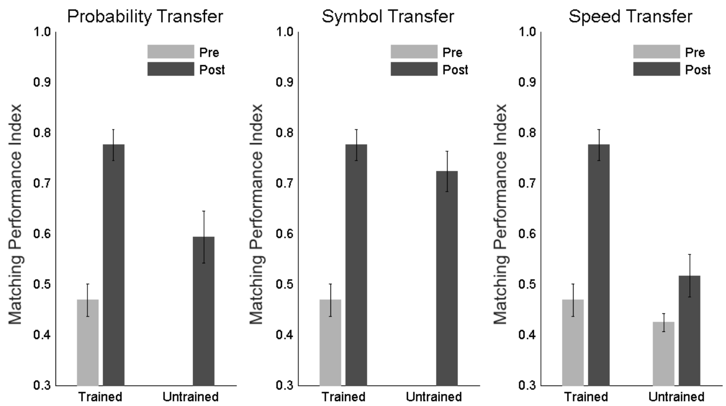
Improvement from the pre-training test to the untrained transfer test occurred (mixed two-way ANOVA, significant main effect of Training, $F(1,53) = 147.59$, $p < .001$), showing generalisation occurred from training to the symbol transfer condition. There was not an interaction effect (Training x Group, $F(3,53) = 2.58$, $p = .06$) nor a main effect of Group, $F(3,53) = .08$, $p =$



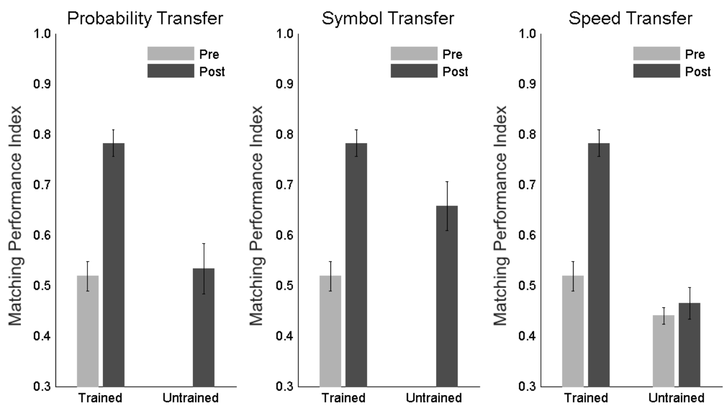
(A) Main



(B) Feedback



(C) Maximization Feedback



(D) No Feedback

FIGURE 3.7: Transfer test performance for participants from the Main (n = 15), Feedback (n = 17), Maximization Feedback (n = 14), and No Feedback (n = 11) groups. Higher matching performance index values indicate more transfer.

.97, indicating that performance did not significantly differ between groups (Figure 3.7).

Speed Transfer Test

Improvement from the pre-training test to the untrained transfer test did not occur (mixed two-way ANOVA, no significant main effect of Training, $F(1,53) = 3.81$, $p = .056$), showing subjects did not generalise from training to the speed transfer condition. There was not an interaction effect (Training x Group, $F(3,53) = 2.56$, $p = .07$) nor a main effect of Group, $F(3,53) = .13$, $p = .94$, indicating that performance did not significantly differ between groups (Figure 3.7).

Direct Transfer Comparisons

This information is the same as presented in (Figure 3.7), but shows a direct comparison of the transfer conditions between groups.

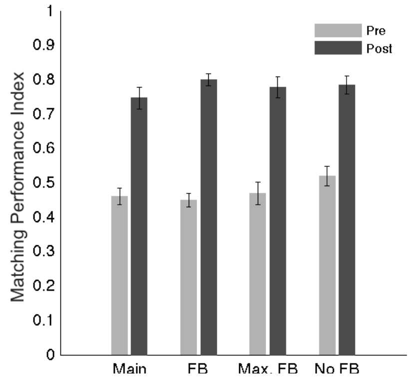
Mean performance in the probability transfer test was not significantly different across groups (one-way ANOVA, $F(3,53) = .87$, $p = .46$) (Figure 3.8(b)).

Mean performance in the symbol transfer test was not significantly different across groups (one-way ANOVA, $F(3,53)= .85$, $p = .47$) (Figure 3.8(c)).

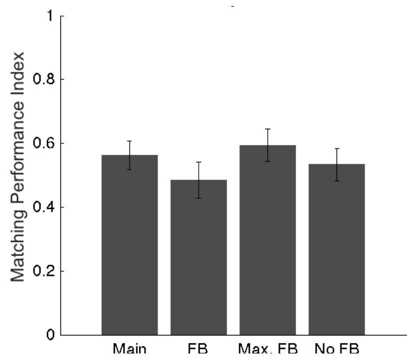
The speed transfer test did not show an interaction effect between Session x Group (mixed two-way ANOVA, $F(3,53) = 1.71$, $p = .18$) nor was there a main effect of Group ($F(3,53) = .49$, $p = .69$), indicating there were no significant differences between groups. However, there was a main effect of Session ($F(1,53) = 21.92$, $p < .001$) indicating some improvement on the transfer task from pre-training transfer test to post-training transfer test (Figure 3.8(d)).

3.4 Discussion

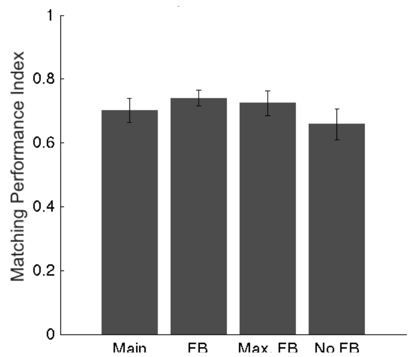
As predicted, learning profiles across groups were similar: differences in test performance, learning rate, and transition times did not reach significance. Probability and symbol transfer test performance was improved from pre-training test performance, demonstrating generalisation. Speed transfer test performance improved between the speed pretest and speed posttest, though perhaps the difficulty of this transfer test prevented strong generalisation. As expected, there were no significant differences between groups with regards to transfer. Most interestingly, strategy profiles were significantly different across groups and while Post-hoc Tukey HSD tests showed no significant differences between individual groups, these results narrowly missed significance. This pattern of results indicates that subjects



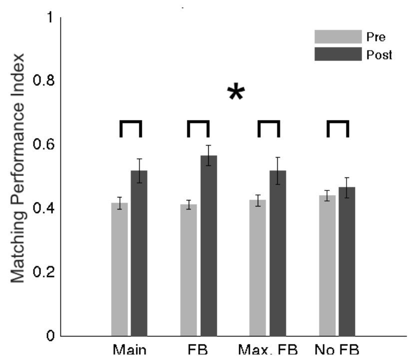
(A) Pre-training and post-training performance for comparison



(B) Probability Transfer



(C) Symbol Transfer



(D) Speed Transfer

FIGURE 3.8: Behavioural test and transfer test performance for participants from the Main (n = 15), Feedback (n = 17), Maximization Feedback (n = 14), and No Feedback (n = 11) groups. Higher matching performance index values indicate more transfer.

in the Feedback group maximized much more than subjects in the Main or No Feedback groups.

As expected from the literature, subjects in the Feedback and Maximization Feedback groups tended to maximize more than subjects in the No Feedback groups (Figure 3.6). What was more unexpected was that the subjects in the Main group had the same ICD and ICD end scores as those in the No Feedback group, whereas any sort of feedback at all (like the block feedback present in the Main condition) was expected to encourage maximization behaviour, and the directed matching feedback was expected to encourage matching behavior. Matching and maximizing strategies both increase ICD and ICD end scores, so subjects in the Main group should have had higher scores than those in the No Feedback condition. This pattern of results could be due to the complication of the paradigm, the hypothesis that directed block feedback does not have a strong effect, or the infrequency of directed block feedback. In the Gao and Corter (2015) study, feedback in the sequence condition was given every 4 trials, whereas in this study feedback was only given every 60 trials.

Another result emerged that was not as predicted: subjects in the Maximizing Feedback group maximized less than subjects in the Feedback group. This finding did not reach significance, but seems to speak to the role of any kind of trial-by-trial feedback in encouraging maximization, directed or not. Future studies should examine whether feedback in any form encourages maximization behavior and the unpredicted directionality observed here was just noise, or if accuracy-based feedback specifically encourages maximization more than maximization-based feedback.

3.4.1 Directed but conflicting feedback: limitations of the study

Subjects have the ability to adjust their maximization and matching strategies in tasks where either one or the other is optimal (Schulze, Ravenzwaaij, and Newell, 2015), and in this study there was a major confound: the "optimal" strategy, as defined by directed feedback, rewarded various and conflicting strategies. In the Feedback task, maximization was implicitly rewarded in the trial-by-trial feedback (in that maximization would produce the highest number of correct answers) but matching was rewarded in the block feedback. In the Maximization Feedback task, maximization was explicitly rewarded in the trial-by-trial feedback, but matching was again rewarded in the block feedback. The fact that the Maximizing Feedback subjects had fewer maximizers than the Feedback subjects hints at the complexity of different types of interacting and conflicting feedback (see Kluger and DeNisi (1996)), and makes it difficult to draw conclusions about the effects of directed feedback in this paradigm. Future studies could employ a more controlled version of this task where task demands were not contradictory: where subjects were singularly encouraged to match or maximize by feedback that incentivised matching, accuracy, and maximization. This experiment would help to isolate which components of feedback subjects were most sensitive to.

3.4.2 What are subjects learning?

Regardless of its cause, the fact that feedback changed subject behaviour prompts the question of whether feedback solely encourages subjects to respond differently or whether it affects the information they process. While it was hoped that feedback would affect subjects' learning, there was only one response measure in this paradigm: which symbol was predicted. This response measure offered higher precision than those of previous statistical learning paradigms, but it too is limited in answering the question of what subjects are learning.

Consider two learners, one who is responding in a maximizing manner and another who is responding in a matching manner. One could assume that both subjects have learned the same statistical contingencies— those required for matching— but that the maximizer is ignoring that knowledge. This would make their responses a function of decision theory; both subjects have learned the same things. On the other hand, perhaps the maximizer discovered the maximizing strategy fairly early on, and did not bother to learn the lower-probability contingencies that the matcher has learned. In this case, the maximizer has stored far less knowledge than the matcher and their responses reflect the knowledge they are representing. Both of these possibilities are identical given the responses, but one set of responses reflects decision theory and the other reflects knowledge representation. This interaction of decision-making and targeted learning illustrates the puzzle of learning, and limits what can be concluded about either knowledge representation or decision-making independently.

3.4.3 Individual differences and future directions

Finally, there are some trends that can be captured via the response measure but are not easily presented because they represent individual differences. Consider a maximizing subject who had started learning all of the contingencies, but forgot them as soon as she/he settled on the maximization strategy; people tend to remember irrelevant distractors less when they are in structured sequences compared to when they are in random sequences (Otsuka and Saiki, 2016) and the maximizer could have decided the lower-probability contingencies were "irrelevant" within the structure they had established. This switch from matcher to maximizer is detectable within the paradigm. However, this sort of switching behaviour is not captured in summary statistics, and the border is nebulous between maximizing, matching, and having a first-order model understanding but an incorrect grasp of context contingencies. All of the summary metrics used in this study— matching performance index, learning rate, ICD— are continuous, which allows individual variability to be identified but not easily classified. These continuous measures allow interesting correlations to be performed (see Chapter 6) and open questions to future investigation but are difficult to summarize.

3.4.4 Conclusions

In this set of experimental groups, we asked whether people changed their strategies when exposed to situations with different desired outcomes. Subjects did engage the prediction task differently in response to feedback, maximizing more even while learning proceeded at similar rates across groups. Further work remains to be done investigating the effects of directed feedback, in addressing the question of behavioural responses compared to learning, and in capturing individual differences in strategies. This work would further contribute to the goal of understanding how people situationally adjust their strategies to direct their statistical learning.

Chapter 4

Temporal Jitter

4.1 Introduction

We are interested in the strategies that people use to learn information. One of the fundamental principles of statistical learning and learning in general is that items can be associated with each other when they are close together in time. When these temporal relationships are disrupted, do peoples' learning strategies change? We sought to determine whether subjects changed their strategies when they were faced with stimuli presented with temporal jitter. We presented regularly-presented stimuli, stimuli with some temporal jitter, and stimuli that were presented with maximal irregularity. We hypothesized that changing the temporal relationships between items would cause subjects to adopt different strategies. Specifically, we expected that subjects would perform more poorly with irregularly-presented stimuli and fail to adopt optimal strategies.

4.1.1 Temporal jitter decreases predictive performance

Regular presentation of material has been shown to increase performance. For example, the concept of "temporal expectancy" describes the anticipation from seeing a cue associated with a target and then seeing a target appear. This cue can be an explicit and independent cue, or an expectation built up by seeing previous items appear in a regular rhythm. The interval between the cue and the target governs reaction times (Woodrow, 1914), and people tend to respond faster in detection and discrimination tasks as the variability of this interval decreases (reviewed in Niemi and Näätänen (1981), e.g. Correa, Lupiáñez, et al. (2004) and Correa, Lupianez, and Tudela (2006)). There is a rich literature indicating that temporal expectation facilitates performance; see Nobre, Correa, and Coull (2007) for a review. On the explicit cue side, temporal expectancies decrease perceptual discrimination thresholds (Lasley and Cohn, 1981; Westheimer and Ley, 1996; Rolke and Hofmann, 2007). On the rhythmic presentation side, perceptual judgments (e.g. of time intervals) improve when items appear at their expected time point in a rhythmic sequence (Jones and M. Boltz, 1989; M.G. Boltz, 1993; Large and Jones, 1999; Barnes and Jones, 2000; Jones, Moynihan, et al., 2002; Guo et al., 2004; Doherty et al., 2005).

For a detailed example, in Jones, Moynihan, et al. (2002) subjects were better at judging pitch (the main task) when there were regular (as opposed to irregular) temporal background auditory cues leading up to the pitch judgment, even though subjects were told that the background cues were irrelevant and to ignore them. However, these background cues still influenced subjects' performance and anticipatory attention. The authors report that subjects listening to the irregular background cues had flatter expectancy profiles compared to subjects who were attentionally primed for the correct time that the stimulus would appear. In another interesting example, Olson and Chun (2001) observed cuing effects with a temporal sequence that was irregular in time but predictably irregular (i.e. stimuli might appear for 80ms, 666ms, 1066ms, 80ms and then loop again). Predictable events can generate temporal expectancies even when complex.

There are neural markers of attention in time, including neural oscillations that align with temporal expectations (e.g. Praamstra et al. (2006) and Rohenkohl and Nobre (2011), see Arnal and Giraud (2012) for a review) and predict perceptual performance (Busch, Dubois, and VanRullen, 2009; Cravo et al., 2013). Regularly- versus irregularly-presented stimuli generate measurably different patterns of neural activity (Zanto, Snyder, and Large, 2006; Lange, 2010; Schwartze et al., 2011).

The jitter used in the current set of experimental groups derailed temporal expectancy because it was added stochastically to sequences to prevent predictability. Based on the body of work outlined here, it would be reasonable to expect that temporal jitter would decrease performance and result in the adoption of strategies that produced less desirable outcomes.

4.1.2 Cognitive load inhibits statistical learning

The fact that violation of temporal expectancies results in worse performance suggests that the surprise from violations distracts attention away from the main task. Given the idea of attentional distraction, a useful framework in thinking about the problem could be cognitive load. It could reasonably be expected that temporal jitter would increase cognitive load on subjects, and cognitive load has been shown to inhibit statistical learning. With regards to artificial language studies, in Saffran, Newport, Aslin, et al. (1997) compared to Saffran, Newport, and Aslin (1996), adults performed poorly (59% accuracy) on a language learning task when they completed a distractor colouring task compared to when they did not have a distractor task (76% accuracy). T. Fernandes, Kolinsky, and Ventura (2010) found that cognitive load decreased performance: subjects in a low-cognitive load condition learned both high- and low- transitional probability words, but subjects in the high-cognitive load condition learned only the high-transitional probability words. Toro, Sinnett, and Soto-Faraco (2005) conducted a similar task with a similar finding: subjects were not able to segment words when they were simultaneously engaging in a high-load distractor task. High cognitive load has also been shown to impair performance in paradigms like the serial reaction time task (Rowland and Shanks, 2006), and other instances are summarized in Perruchet and Pacton (2006). In Baker et al. (2014), fewer subjects could perform their temporal sequence task under high cognitive load. These results suggest that temporal jitter could produce worse performance and poorer-performing strategies.

4.1.3 Jitter as increasing chunking

Finally, one of the most important considerations for how temporal jitter might affect learning is the hypothetical formation of chunks. Before we address chunking, we must ask: how do people learn temporal relationships? In the artificial grammar learning literature, the question of "what are people learning?" has been extensively explored. Originally it was argued that subjects were extracting rules from artificial grammars, and applying those rules to determine grammaticality at test (Reber, 1967). Since then arguments have been raised that what people are learning is actually simpler: some form of similarity mapping between training and test exemplars (e.g. Vokey and Brooks (1992)). In addition, since the artificial grammar learning stimuli are composed of relatively short sequences, it is commonly accepted that some aspect of what subjects are learning is chunks (e.g. Lotz and Kinder (2006), see Perruchet and Pacton (2006)), often bigrams or trigrams.

The idea of chunking, or learning short sequences of items, is an intuitive way of storing and remembering information. Temporal jitter could be responsible for forming chunks: placing certain items closer together temporally and others further apart. Dividing long sequences into chunks could result in faster learning and strategies that result in better performance, especially if these chunks align with the sequence structure of the task. On the other hand, if these chunks do not align with the sequence structure of the task— if they group items in trigrams, for example, when the true structure is a quartet— they may slow learning as subjects become less flexible in what rules they learn. The question of what effect temporal jitter will have on chunking is further complicated by the fact that temporal jitter in this task is randomly applied, so that it is impossible to know beforehand the structure of chunks that will emerge and the chunks change on every trial.

4.1.4 Experiment

In this section we set out to ask whether subjects' strategies would change when the temporal relationships between items were disturbed. Since temporal relationships drive almost all statistical learning (excepting spatial visual statistical learning), this manipulation was expected to produce reliable differences in behaviour. Based on the detrimental effects of irregular presentation in the literature, it was expected that subjects would perform more poorly when temporal jitter was added to stimulus presentation. Additional evidence for the hypothesis that temporal jitter would reduce performance comes from the literature on the effects of cognitive load assuming that cognitive load increases as temporal jitter is added.

Independent groups of subjects observed stimuli with three degrees of temporal jitter. The first group, "Main", was previously described, had no jitter, and served as the control group. The stimuli in the "Jitter" and "Augmented Jitter" groups both had randomly-allocated temporal jitter, but in the Augmented Jitter group jitter was applied in intervals that spanned a wider time range and were more finely sampled. Specifically, in the Main group the inter-stimulus interval (ISI) between symbols was 400ms, in the Jitter group ISI ranged from 200-600ms in intervals of 100ms, and in the Augmented Jitter group ISI ranged from 200-700ms in intervals of 20ms.

It was hypothesized that subjects would perform progressively more poorly with increasing amount of jitter. Specifically, average matching performance index was expected to be the highest in the Main group and the lowest in the Augmented Jitter group. Learning rate was expected to follow a similar trend: highest scores in the Main group while the worst performance and lowest scores would be found in the Augmented Jitter group. Transition times were expected to be lowest in the Main group, indicating fastest learning, and highest in the Augmented Jitter group. With regards to strategies, subjects in the Main group were expected to have high ICD and ICD end values showing successful acquisition of maximizing and matching strategies compared to the Augmented Jitter group, which were expected to have the lowest and possibly negative ICD and ICD end values showing a failure to acquire successful matching or maximizing strategies. With regards to transfer tests— probability, symbols, and speed— subjects were again expected to have progressively poorer generalisation as temporal jitter increased. Scores were hypothesized to be highest in the Main group, lower in the Jitter group, and lowest in the Augmented Jitter group.

TABLE 4.1: Jitter Experimental Groups Summary. Groups are in the following order: Main, Jitter (Jit.), Augmented Jitter (Aug. Jit.).

Groups:	**Main**	**Jit.**	**Aug. Jit.**
Temporal Jitter (up to %)	0	40	60
Markov Model	main	main	main
Trial-By-Trial Feedback	n/a	n/a	n/a
Block Feedback	yes	yes	yes

4.2 Methods

Three experimental groups were compared: "Main", "Jitter", and "Augmented Jitter". Different amounts of temporal jitter were added to the presented sequences for the Jitter group (up to 40% of possible jitter) and Augmented Jitter group (up to 60% of possible jitter). Specifically, subjects in the Main, Jitter, and Augmented Jitter groups saw stimuli for 100ms. However, in the Main Group, the inter-stimulus interval (ISI) was always 400ms, while this ISI was randomly varied for the Jitter and Augmented Jitter groups. In the Jitter group, ISI blocks had mean 400ms but were uniformly and randomly distributed across the values 200/300/400/500/600ms (Jitter). In the Augmented Jitter group, ISI blocks had mean 400ms but were uniformly and randomly distributed across the values 100/120/.../680/700ms. All groups had block feedback (encouraging maximization) (Table 5.1).

Subjects were trained over four days, and on the last day did a testing sequence (test block, random block, test block), the probability transfer test (the high- and low-probability statistical contingencies were switched), the symbol transfer test (symbols were replaced by a new set of symbols), and then the speed transfer test (items were presented more quickly).

4.3 Results

4.3.1 Learning Profile

Matching Performance Index

Subjects had similar learning profiles across the different levels of jitter. Matching performance index patterns across training were similar across groups (Figure 4.1). Mean test scores (pre- and post-) were not significantly different across jitter groups (mixed two-way ANOVA, Session x Group, $F(2,41) = .27$, $p = .76$). There was not a main effect of Group ($F(2,41) = 1.66$, $p = .20$). There was a significant main effect of Session, $F(1,41) = 183.31$, $p <$

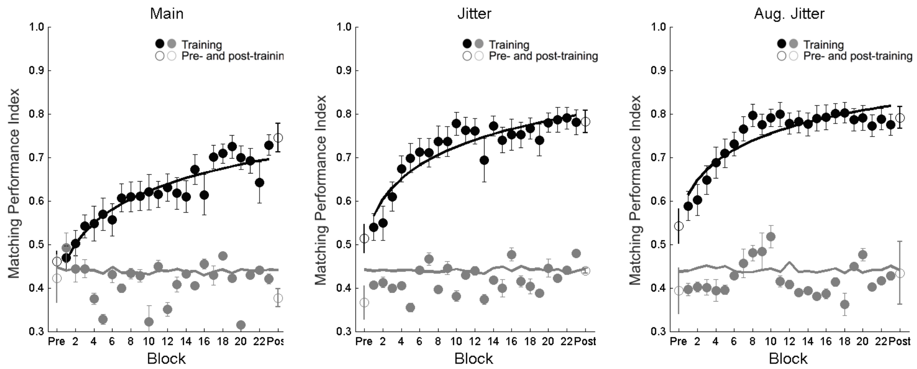


FIGURE 4.1: Behavioural performance. Matching performance index for participants from the Main (n = 15, weak learners: n = 3), Jitter (n = 14, weak learners: n = 4), and Augmented Jitter (n = 15, weak learners: = 3) groups across training (solid circles), the pre-training test (open circles) and the post-training test (open circles). Higher scores on the matching performance index indicates more matching behaviour. Participants completed the task over five days. Data is fitted for participants who improved during training (black circles). Data is also shown for participants that did not improve during training (grey circles). Random guess baseline is indicated by a solid grey line across blocks. Error bars indicate standard error of the mean.

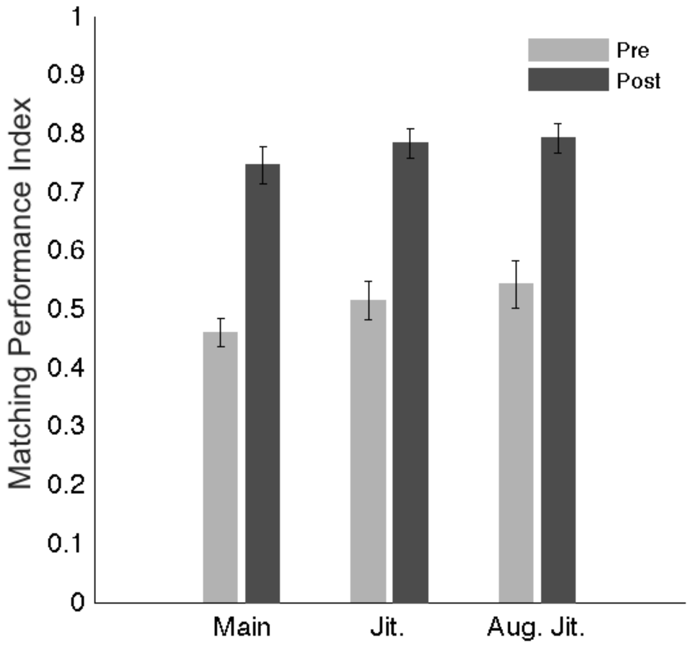


FIGURE 4.2: Behavioural test performance. Matching performance index for participants from the Main (n = 15), Jitter (n = 14), and Augmented Jitter (n=15) participants for pre-training performance and post-training performance. Higher scores on the matching performance index indicates more matching behaviour.

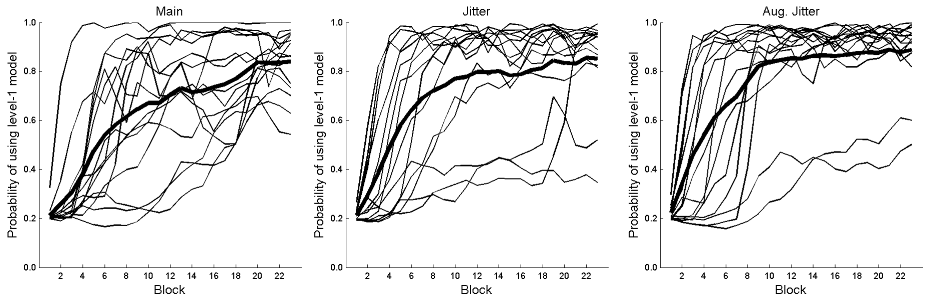


FIGURE 4.3: Learning curves. Mixture coefficient weights for level-1 compared to level-0 model for participants from individual Main (n = 15), Jitter (n = 14), and Augmented Jitter (n=15) groups. The average curve is shown as a thicker line. Weights closer to 1 indicate that the subject is more likely to be making predictions based on a Markov level-1 model; weights closer to 0 indicate that the subject is more likely to be making predictions based on a Markov level-0 model.

.001 (Figure 4.2), which indicated that subjects improved from pre-training to post-training.

Learning Indices

Learning profiles appeared to show improved learning (higher learning rates, lower transition times) in groups with more temporal jitter (Figure 4.3). However, mean learning rates were not significantly different across groups (one-way ANOVA, $F(2,41) = 2.32$, $p = .11$), so subjects did not quantitatively learn faster in different groups (Figure 4.4). Mean transition times were not significantly different across groups (one-way ANOVA, $F(2,41) = 1.37$, $p = .27$), so subjects did not quantitatively learn earlier in different groups (Figure 4.4).

To test for a linear trend between groups, groups were assigned a continuous measure for jitter: the stimuli for the Main group had up to 0% jitter and so was assigned 0, the stimuli for the Jitter group had up to 40% jitter so was assigned 40, and the stimuli for the Augmented Jitter group had up to 60% jitter and so was assigned 60. Linear regression with learning rate as the dependent variable, and a constant and the degree of jitter as independent variables, was run. Learning rate across groups was not described by a linear relationship ($R = .24$, $F(1,42) = 2.50$, $p = .12$). Linear regression with transition time as the dependent variable, and a constant and the degree of jitter as independent variables, was run. Transition time across groups was not described by a linear relationship ($R = .25$, $F(1,42) = 2.80$, $p = .10$).

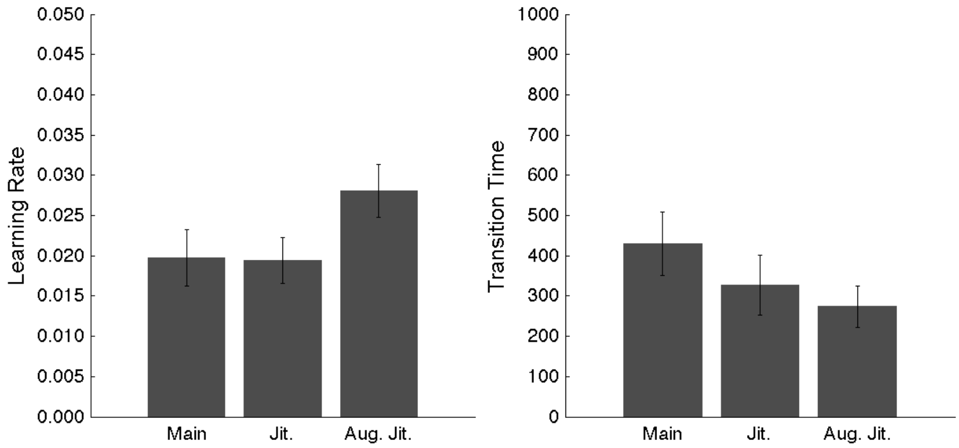


FIGURE 4.4: Mean learning indices for participants from Main (n = 15), Jitter (n = 14), and Augmented Jitter (n=15) groups. Learning rate is the slope of the learning sigmoid curve; transition time is the time point at which y-axis of the learning curves is equal to .5 (the weights for level-1 and level-0 Markov models are both equal to .5). Higher learning rates indicate a faster rate of learning; lower transition time values indicate earlier learning.

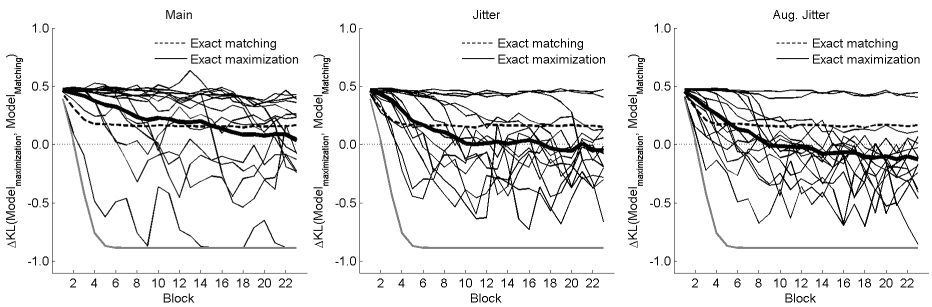


FIGURE 4.5: Strategy choice. ΔKL divergence between model matching and model maximization strategies for participants from the Main (n = 15), Jitter (n = 14), and Augmented Jitter (n=15) groups. KL divergence is a measure of how different the probability distributions that the subject is using to make predictions are from the probability distributions of predictions made using a perfect maximization strategy and a perfect matching strategy. The average curve is shown as a thicker line. Exact matching (dashed line) and maximization models (solid grey line) are plotted. More negative KL scores indicate a strategy closer to maximization.

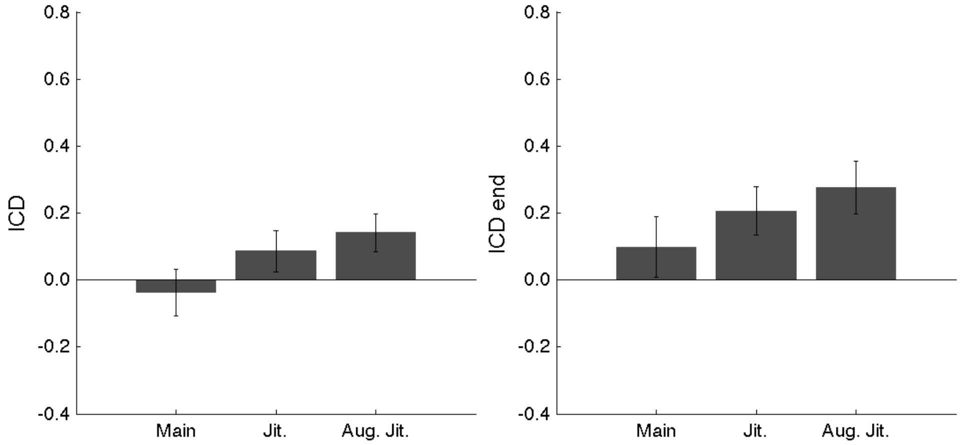


FIGURE 4.6: Mean strategy indices for participants from the Main (n = 15), Jitter (n = 14), and Augmented Jitter (n=15) groups. ICD and ICD end measure the signed area between the subjects' strategy curve and predictions made using a perfect matching strategy. Higher ICD and ICD end values indicate a strategy closer to maximization. ICD and ICD end values equal to zero indicate a perfect matching strategy.

Strategy

Strategy profiles appeared to show maximization increasing across groups as jitter was increased (Figure 4.5). However, ICD scores were not significantly different across groups (one-way ANOVA, F(2,41) = 2.12, p = .13) (Figure 4.6). ICD end scores were not significantly different across groups (one-way ANOVA, F(2,41) = 1.25, p = .30) (Figure 4.6).

Because there appeared to be a trend in which temporal jitter encouraged maximization, we tested for a linear trend between experimental groups. Groups were assigned a continuous measure for jitter: the stimuli for the Main group had up to 0% jitter and so was assigned 0, the stimuli for the Jitter group had up to 40% jitter so was assigned 40, and the stimuli for the Augmented Jitter group had up to 60% jitter and so was assigned 60. Linear regression with ICD as the dependent variable, and a constant and the degree of jitter as independent variables, was run. ICD across groups was described by a linear relationship (R = .31, F(1,42) = 4.34, p = .043). This finding indicated that subjects in the experimental groups with more jitter followed strategies closer to maximization. Linear regression with ICD end as the dependent variable, and a constant and the degree of jitter as independent variables, was run. ICD end across groups was not described by a linear relationship (R = .24, F(1,42) = 2.55, p = .12), though it followed the same trend as ICD across groups.

4.3.2 Transfer

Probability Transfer Test

In the probability transfer test the high- and low-probability statistical contingencies were exchanged. Improvement from the pre-training test to the untrained transfer test did not occur (mixed two-way ANOVA, no main effect of Training, $F(1,41) = .90$, $p = .35$), indicating no generalisation occurred from training to the probability transfer condition. There was not an interaction effect of Training x Group, $F(2,41) = 1.67$, $p = .20$, nor a main effect of Group, $F(2,41)= 1.00$, $p = .38$, indicating that performance did not significantly differ between groups (Figure 4.7).

Symbol Transfer Test

Improvement from the pre-training test to the untrained transfer test occurred (mixed two-way ANOVA, significant main effect of Training, $F(1,41) = 66.77$, $p < .001$), showing generalisation occurred from training to the symbol transfer condition. There was not an interaction effect (Training x Group, $F(2,41) = .70$, $p = .50$) nor a main effect of Group, $F(2,41) = .79$, $p = .46$, indicating that performance did not significantly differ between groups (Figure 4.7).

Speed Transfer Test

Improvement from the pre-training test to the untrained transfer test did not occur (mixed two-way ANOVA, no main effect of Training, $F(1,41) = .31$, $p = .58$), indicating no generalisation occurred from training to the speed transfer condition. There was not an interaction effect (Training x Group, $F(2,41) = 3.12$, $p = .06$) nor a main effect of Group, $F(2,41) = .24$, $p = .79$, indicating that performance did not significantly differ between groups (Figure 4.7).

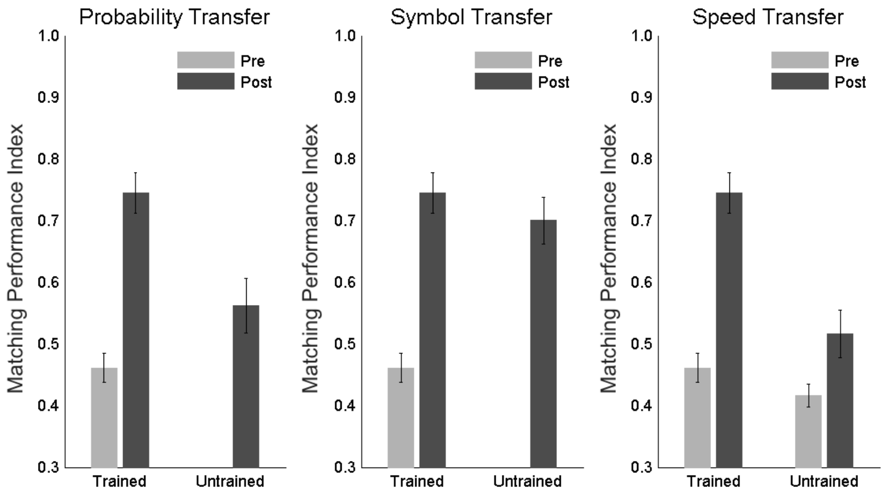
Direct Transfer Comparisons

This information is the same as presented in (Figure 4.7), but shows a direct comparison of the transfer conditions between groups.

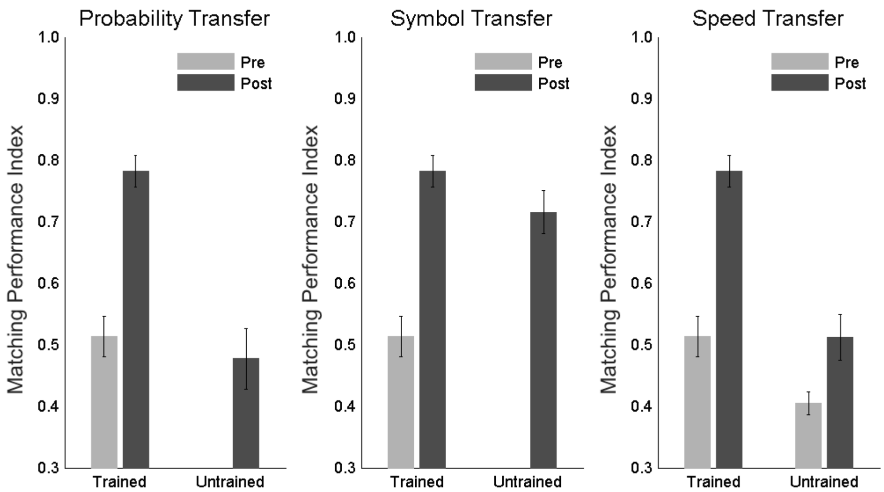
Mean performance in the probability transfer test was not significantly different across groups (one-way ANOVA, $F(2,41) = 1.15$, $p = .33$) (Figure 4.8(b)).

Mean performance in the symbol transfer test was not significantly different across groups (one-way ANOVA, $F(2,41)= .05$, $p = .96$) (Figure 4.8(c)).

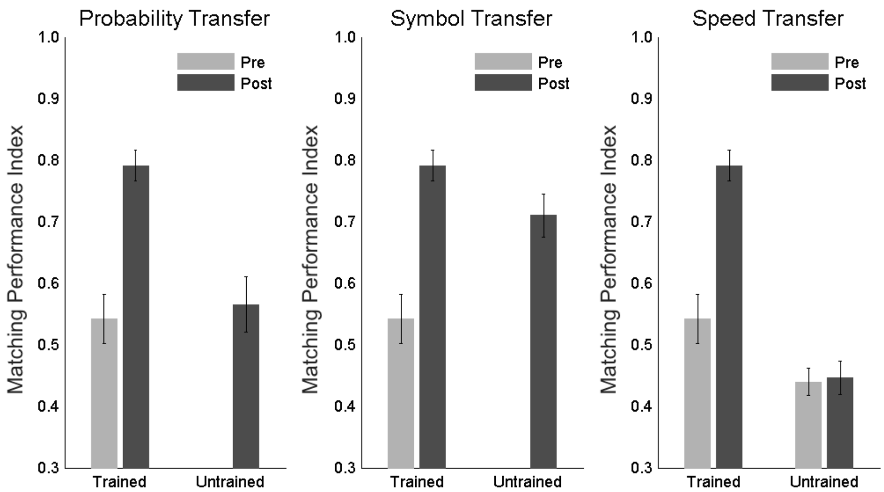
The speed transfer test did not show an interaction effect between Session x Group (mixed two-way ANOVA, $F(2,41) = 2.54$, $p = .09$) nor a main effect of Group ($F(2,41) = .31$, $p = .74$), indicating there were no significant



(A) Main

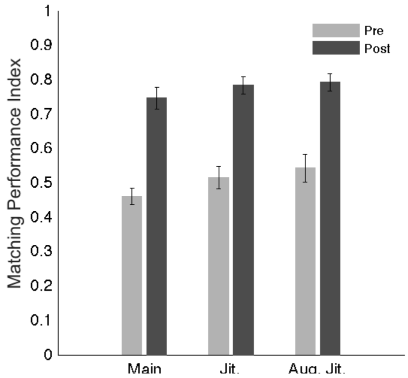


(B) Jitter

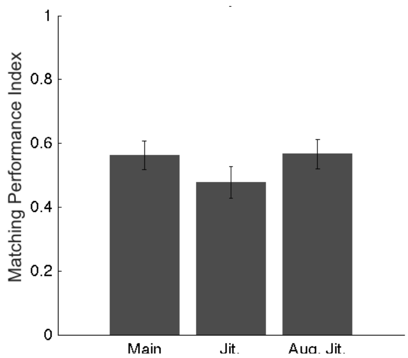


(C) Augmented Jitter

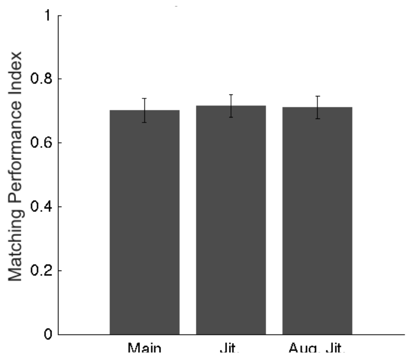
FIGURE 4.7: Transfer test performance for participants from the Main (n = 15), Jitter (n = 14), and Augmented Jitter (n=15) groups. Higher matching performance index values indicate more transfer.



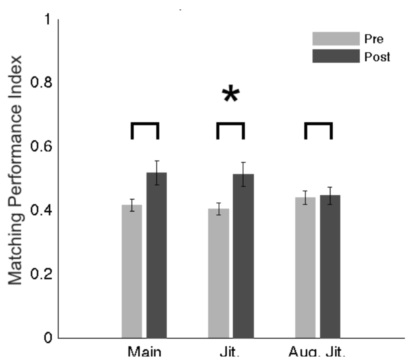
(A) Pre-training and post-training performance for comparison



(B) Probability Transfer



(C) Symbol Transfer



(D) Speed Transfer

FIGURE 4.8: Behavioural test and transfer test performance for participants from the Main (n = 15), Jitter (n = 14), and Augmented Jitter (n=15) groups. Higher matching performance index values indicate more transfer.

differences between groups. However, there was a main effect of Session ($F(1,41) = 12.25$, $p < .001$) indicating some improvement on the transfer task from pre-training transfer test to post-training transfer test (Figure 4.8(d)).

4.4 Discussion

In this section we asked whether subjects would change their strategies when a fundamental element of learning— temporal relationships— was altered by introducing temporal jitter. We hypothesized that subjects would have poorer performance in conditions of more temporal jitter. Specifically, we expected lower performance indices, lower learning rates, and higher transition times in the Augmented Jitter group, middling performance for the Jitter group, and opposite performance for the Main group. In terms of strategy, we expected lower ICD and ICD end values in the Augmented Jitter group and higher ICD and ICD end values in the Main group showing acquisition of successful maximizing and matching strategies. In terms of generalisation, scores were expected to be lowest in the Augmented Jitter group and highest in the Main group.

Instead, we observed that learning profiles across experimental groups were similar: differences in test performance, learning rate, and transition times did not reach significance, and follow-up statistical tests with learning rate confirmed this finding. Learning rates and transition times did not follow a linear relationship across groups. Strategy profiles also appeared somewhat similar: differences in ICD and ICD end did not reach significance, and follow-up tests with ICD and ICD end confirmed this finding. Symbol transfer test performance was improved from pre-training test performance, demonstrating generalisation. However, there were no significant differences between groups with regards to transfer. Another measure of poorer performance in groups with temporal jitter could have been an increased number of weak learners, but these numbers were similar across groups (the number of weak learners in the Main group was 3/18, Jitter was 4/18, and Augmented Jitter was 3/18.)

However, the most interesting result was that while strategy differences across groups were not significant, there was a qualitative trend whereby ICD increased as temporal jitter increased. This finding was supported quantitatively as ICD was described by a positive linear trend in groups with increasing jitter, meaning that subjects who were exposed to more jitter adopted strategies closer to maximizing. The same positive linear trend was shown for ICD end, though it did not reach significance. These trends were also observed in the learning profiles: a close-to-significant positive trend in learning rate (higher learning rates for the groups with more jitter), and a close-to-significant negative trend in transition time (lower transition times for the groups with more jitter). Given the small number of subjects per group and individual variability, it is remarkable that these trends— all in one direction— emerged in all of the response measures. The results are consistent and opposite from those predicted: subjects in groups with more temporal jitter performed *better* than subjects in groups with less temporal jitter.

4.4.1 Why did jitter produce stronger learners? Hardness hypothesis

It was predicted that jitter would make subjects weak learners; however, the pattern of results across learning rate, transition time, and strategy measures all support the conclusion that irregularly-presented stimuli improves performance. Previous studies have indicated that irregularly-presented stimuli often decrease performance. Why would this unusual pattern of results emerge?

One hypothesis harks back to the "maximization as a dumb strategy" perspective discussed in Chapter 1. This argument states that maximizing is easier than matching, and in the experiment described here, it is certainly easier to remember one relationship for each context— "C" usually follows "B"— rather than two relationships for each context— "A" follows "B" 20% of the time and "C" follows "B" 80% of the time). It is possible that when faced with jitter, subjects are placed under increased cognitive load and are overwhelmed. Flustered, they may have decided to settle on the easier maximization strategy, which would explain this pattern of results: higher ICD and ICD end values, meaning more maximizers, and a faster learning rate and earlier transition time as subjects forwent trying to learn a complicated matching strategy for learning a simpler maximizing strategy. This explanation would fit into the context of the literature. This pattern of results would then provide a powerful example of how increasing temporal jitter forces subjects to progressively adopt strategies that optimize their performance in complex learning problems.

Further testing of this "hardness" hypothesis is necessary. As a control condition, more traditional cognitive load should be applied to a separate group of participants to observe if the same pattern of responses— more cognitive load, more maximization and faster learning— occurs. An interesting interaction effect could also occur if additional and independent cognitive load were applied to all of this study's groups. One would expect more maximization to emerge among the low temporal jitter groups, but that additional maximization would not occur among the high temporal jitter groups. If applying additional cognitive load resulted in the same increase in maximization across all groups, this would suggest that subjects in the Augmented Jitter condition had not reached their maximum or optimal cognitive load, and should be placed under more pressure to optimize performance.

4.4.2 Why were effects weak? People as probabilistically rational learners

The pattern of results observed— that subjects exposed to more temporal jitter performed better— was consistent across measures but weak. One possible reason for these weak results could be the effects of small sample sizes and high individual variability. Another reason is the small differences between the jitter conditions— in the Jitter condition, ISIs varied randomly between 200 and 600ms, while in the Augmented Jitter condition, ISIs varied randomly between 100 and 700ms. These were relatively small manipulations, and the fact that differences could be observed between groups at all speaks to the power of these interventions.

However, to take another tack, an intriguing explanation for why there were not large differences across experimental groups— groups with increasing amounts of jitter— is the idea that it is rational to ignore uninformative jitter. If temporal jitter is treated as just a "surface feature" of the stimuli, like the spatial position / colour / shape of the symbols, rather than as a fundamental property governing the essential temporal relationships between items, then predictions change. If temporal jitter were just a surface feature of the stimuli, like colour, then temporal jitter would be a completely non-informative feature since it is randomized, and performance would be expected to be roughly the same across all experimental groups.

If jitter is considered an informationally-valueless surface feature, then it makes rational sense for subjects to ignore it and perform equally well with or without it. Aslin and Newport (2012) have overviewed evidence for the idea that people are rational probabilistic learners who base their generalisations on context: people use the reliability / uniqueness of cues to find the right cues for learning structure. In the surface-feature perspective, temporal jitter is a form of noise in this task— it was not informative for predicting the next symbol, so subjects should logically have ignored it. People have widely demonstrated this rational learning ability. For example, in Gerken (2006) and Gerken (2010), infants in an artificial language learning paradigm learned the abstract rule that was the most specific but still reliable given the data. Similarly, Reeder, Newport, and Aslin (2009) and Reeder, Newport, and Aslin (2010) investigated artificial word learning in adults and found that generalisation proceeds in a probabilistically rational way. Turk-Browne, Isola, et al. (2008) has found that subjects are sensitive to the amount of informational covariance between two features, shape and colour, and adjust their behaviour accordingly.

If temporal jitter is considered a surface feature, a human capacity for dismissing it as informationally-irrelevant seems rooted in strong drives for rational thinking and behaviour. If this is the case, then it would not be surprising that subjects would not change their behaviour overmuch when faced with stimuli that were presented in a randomly-irregular way. Subjects would be expected to have similar performance across experimental groups.

This hypothesis needs further testing. Temporal relationships seem too fundamental to reduce to a surface feature, and it seems likely that small sample sizes are the more realistic reason for why the effect sizes across experimental groups were not large. To test the small sample sizes hypothesis, more subjects could be tested. To test the "surface feature" hypothesis, the informational content of temporal jitter could be parametrically increased and compared with the informational content of a more classical surface feature like colour. For example, an irregularly-presented sequence with no relation to the task (e.g. *fast-fast-fast-slow-slow-slow*) would have an information content of zero, but an irregularly-presented sequence predictive in the task (e.g. *fast for A, slow for B, fast for C, slow for D*) would have a higher information content. These levels of information content could be replicated with colour. If, at the same information level, temporal jitter affects performance more than the equivalent colour cue, this describes the degree to which temporal information is weighted as more than a surface feature. This test would provide a quantitative and parametric measure of the relative value of temporal jitter as a fundamental and uniquely important property of how we learn, irrespective of information content.

4.4.3 Conclusions

In this section we asked whether changing temporal relationships would change how subjects learn. Subjects were split into three groups and exposed to different degrees of temporal jitter: some subjects viewed stimuli that were regularly presented, and two other groups viewed stimuli that were presented with temporal jitter to a lesser and greater degree. The relationship between items in time is the most important element of most statistical learning, and so it was expected that varying the stimulus presentation rate would cause subjects to change their strategies for learning. Subjects were expected to perform more poorly when exposed to more temporal jitter.

However, the opposite result occurred. Though many of these results did not pass the $p < .05$ threshold, subjects systematically increased their performance when exposed to temporal jitter. They learned faster and adopted more maximization strategies. One explanation for why this behaviour emerged is the "hardness hypothesis"— subjects were overwhelmed by the difficulty of the tasks with temporal jitter and so adapted the easier maximization strategy. Results might not have been strong because of small sample sizes, high individual variability, or possibly information-content concerns.

This pattern of results revealed that subjects did change their strategies in response to temporal jitter, and if the "hardness hypothesis" holds true, then they changed their strategies in a way consistent with the attentional and cognitive load considerations in the literature. The unpredicted results, however, are intriguing and suggest the need for further study. Why did subjects adopt the maximizing strategy when faced with temporal jitter—was it a conscious decision or unconscious? Could this effect occur with any kind of cognitive load? Does this effect occur only on complex tasks? Future study is necessary to determine exactly how the regularity of stimulus presentation affects the strategies subjects use to learn.

Chapter 5

Structural Contingencies

5.1 Introduction

In this section we probe most deeply into the question of which strategies subjects use to learn. There are two modal strategies subjects use when responding to probabilities: maximization and matching. Both affect what rewards one gains in life— some are more optimal in different situations— so choice of strategy matters. Maximization is easier in most contexts: just find the result that is most often correct, and learn that rule. Matching requires learning an entire probability distribution. When do people choose to use one strategy over another?

We addressed this question by constructing two experimental conditions that were roughly equated in the effort required to maximize, but very different in the effort required to match. We hypothesized that subjects would approach the problem in a graded manner based on difficulty: many subjects would maximize, because they found it easy. However, more subjects would match in the easier matching condition compared to the harder matching condition. Creating these conditions let us ask whether subjects considered the "difficulty" of a strategy before using it, which would seem a computationally rational idea to pursue.

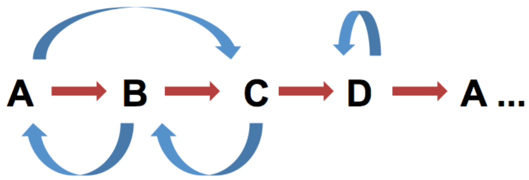
Two experimental groups were compared. In both groups, the high-probability sequence was the same. However, one group had an odd low-probability loop structure (the same control group "Main" introduced previously), while the other group had a low-probability loop that was simply the reverse of the high-probability sequence (the "Different Contingencies") group. In this paradigm, maximizing only requires learning the high-probability loop. Matching requires learning the low-probability statistical contingencies as well. Because the low-probability statistical contingencies were more complicated in the Main group, it was expected that there would be fewer matchers in the Main group.

In this study, both strategies— maximization and matching— were combined into a single continuous measure, ICD and ICD end. A matching strategy resulted in lower strategy values, while a maximizing strategy resulted in higher strategy values. We predicted that there would be similar numbers of maximizers in each group, since this strategy was roughly the same difficulty in both conditions. (The maximization strategy might have been slightly harder to learn in the Main group, because the low-probability statistical contingencies could sometimes restart the subject in a different part of the loop than they were expecting. This unexpected switching did not occur in the Different Contingencies condition, where subjects moved backwards in the loop one step at a time if they were being towed by low-probability contingencies.) We predicted there would be more matchers in the Different Contingencies group due to the ease of learning the low-probability statistical contingencies, and therefore that there might be a few more maximizers in the Main group to even out the number of total maximizers and matchers across groups. Thus, subjects in the Main condition were hypothesized to have higher ICD and ICD end values, because they would have a slightly higher number of maximizers and fewer matchers compared to the Different Contingencies condition. Subjects in the Main condition were expected to have lower matching performance index values for the same reason (fewer matchers).

Subjects in both groups were expected to have roughly the same learning rates and transition times since both matching and maximization strategies increased these values at roughly the same rate. In transfer test results, subjects in the Different Contingencies condition were expected to do much better in the probability transfer test. In the probability transfer test, the high-probability statistical contingencies become the low-probability statistical contingencies and vice versa. It should have been much easier for the Different Contingencies subjects to learn their new "high-probability" loop (since this loop was almost exactly the same as their training loop, but reversed, see Figure 5.1(b)) compared to the Main subjects (who were trying to learn oddly structured "A"→"C"→"B"→"A" / "D"→"D" statistical contingencies, see Figure 5.1(a)). Subjects were hypothesized to perform roughly equally in the symbol transfer and speed transfer tests, since these tests did not unduly rely on differences in strategy.

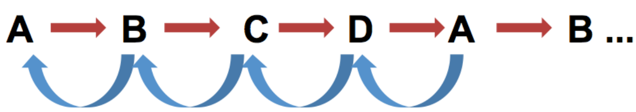
5.2 Methods

Two experimental groups were compared: "Main" and "Different Contingencies". These groups were roughly equated in difficulty for maximizers (they contained the same high-probability loop) but different in difficulty for matchers (the "Main" group had more complex low-probability



Level 1		Target				P (c)
		A	B	C	D	
Context	A		0.8	0.2		0.25
	B	0.2		0.8		0.25
	C		0.2		0.8	0.25
	D	0.8			0.2	0.25

(A) Markov model design for Main, Jitter, Augmented Jitter, Feedback, Maximization Feedback, and No Feedback groups.



Level 1		Target				P (c)
		A	B	C	D	
Context	A		0.8		0.2	0.25
	B	0.2		0.8		0.25
	C		0.2		0.8	0.25
	D	0.8		0.2		0.25

(B) Alternative Markov model design for Different Contingencies group.

FIGURE 5.1: Markov model designs. Red arrows indicate the most likely symbol to appear given the context (80%); blue arrows indicate the less likely symbol (20%). P(c) refers to marginal probabilities. "A", "B", "C", and "D" are replaced by symbols (randomized by participant).

TABLE 5.1: Structural Contingencies Experimental Groups Summary. Abbreviations: "Cont." refers to "Different Contingencies", "alt." refers to the alternate Markov model.

Groups:	**Main**	**Cont.**
Temporal Jitter (up to %)	0	0
Markov Model	main	alt.
Trial-By-Trial Feedback	n/a	n/a
Block Feedback	yes	yes

loops than the "Different Contingencies" group). The "Main" group and the "Different Contingencies" groups differed in the first-order Markov models used to generate the sequences (Figure 5.1). Both designs had block feedback (encouraging matching) and no jitter (Table 5.1).

Subjects were trained over four days, and on the last day did a testing sequence (test block, random block, test block), the probability transfer test (the high- and low-probability statistical contingencies were switched), the symbol transfer test (symbols were replaced by a new set of symbols), and the speed transfer test (items were presented more quickly).

5.3 Results

5.3.1 Learning Profile

Matching Performance Index

Subjects had similar learning profiles across the different structural contingencies groups. Matching performance index patterns across training were similar across groups (Figure 5.2). Mean test scores (pre- and post-) were not significantly different across groups (mixed two-way ANOVA, Session x Group, $F(1,27) = .17$, $p = .68$). There was not a main effect of Group ($F(1,27) = 1.29$, $p = .27$). There was a significant main effect of Session, $F(1,27) = 155.01$, $p < .001$ (Figure 5.3), which indicated that subjects improved from pre-training to post-training.

Learning Indices

Learning profiles appeared similar across groups (Figure 5.4). Mean learning rates were not significantly different across groups (independent samples 2-tailed t-test, $t(27) = -.47$, $p = .64$), so subjects did not learn faster in different groups (Figure 5.5). Mean transition times were not significantly different across groups (independent samples 2-tailed t-test, $t(27) = .81$, $p = .42$), so subjects did not learn earlier in different groups (Figure 5.5).

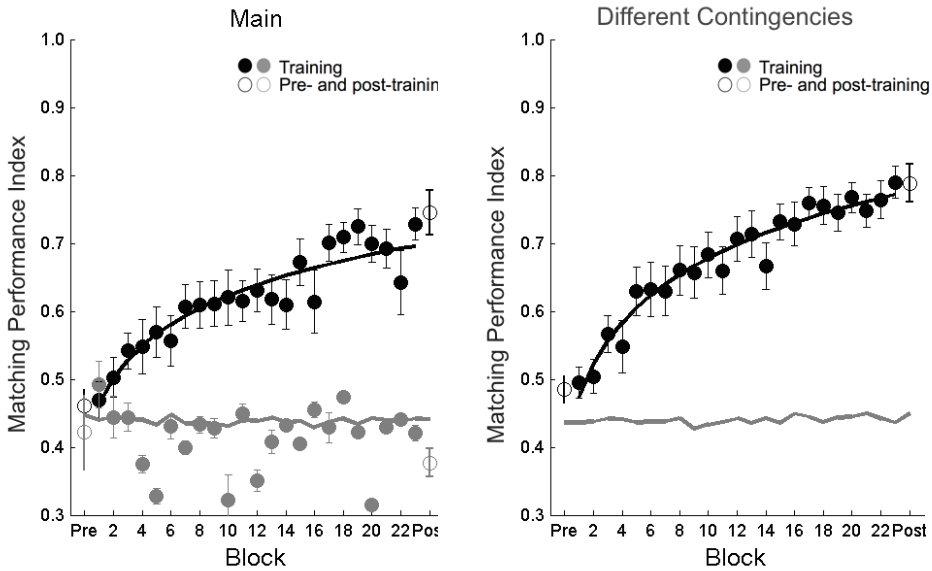


FIGURE 5.2: Behavioural performance. Matching performance index for participants from the Main (n = 15, weak learners: n = 3) and Different Contingencies (n = 14, weak learners: n = 0) groups across training (solid circles), the pre-training test (open circles) and the post-training test (open circles). Higher scores on the matching performance index indicates more matching behaviour. Participants completed the task over five days. Data is fitted for participants who improved during training (black circles). Data is also shown for participants that did not improve during training (grey circles). Random guess baseline is indicated by a solid grey line across blocks. Error bars indicate standard error of the mean.

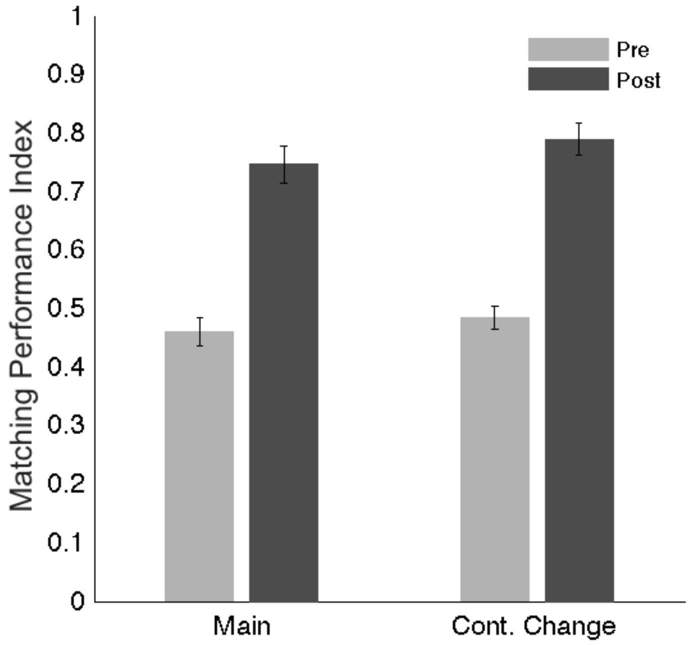


FIGURE 5.3: Behavioural test performance. Matching performance index for participants from the Main (n = 15) and Different Contingencies (n = 14) groups for pre-training performance and post-training performance. Higher scores on the matching performance index indicates more matching behaviour.

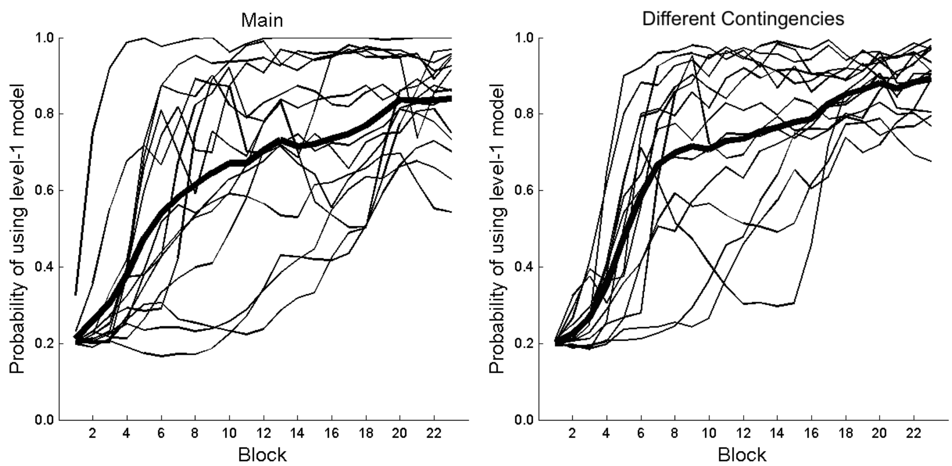


FIGURE 5.4: Learning curves. Mixture coefficient weights for level-1 compared to level-0 model for participants from Main (n = 15) and Different Contingencies (n = 14) groups. The average curve is shown as a thicker line. Weights closer to 1 indicate that the subject is more likely to be making predictions based on a Markov level-1 model; weights closer to 0 indicate that the subject is more likely to be making predictions based on a Markov level-0 model.

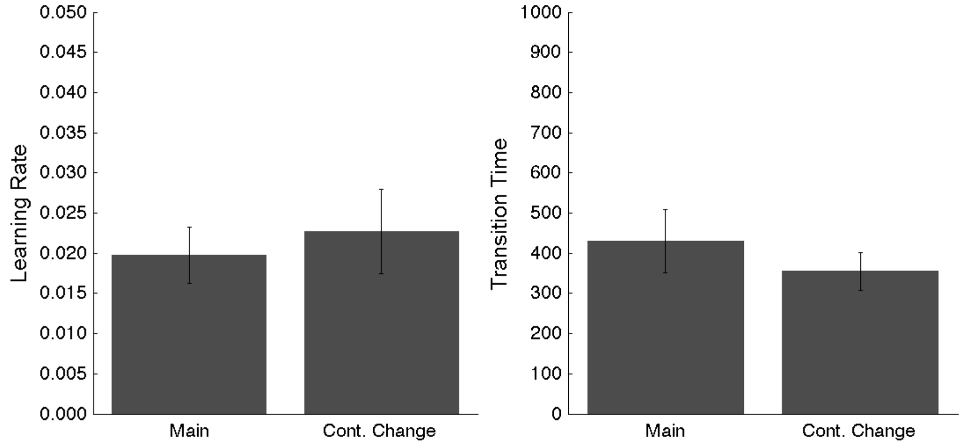


FIGURE 5.5: Mean learning indices for participants from Main (n = 15) and Different Contingencies (n = 14) groups. Learning rate is the slope of the learning sigmoid curve; transition time is the time point at which y-axis of the learning curves is equal to .5 (the weights for level-1 and level-0 Markov models are both equal to .5). Higher learning rates indicate a faster rate of learning; lower transition time values indicate earlier learning.

Strategy

Strategy profiles appeared similar across groups (Figure 5.6). ICD scores were not significantly different across groups (independent samples 2-tailed t-test, $t(27) = .64$, $p = .53$) (Figure 5.7). ICD end scores were not significantly different across groups (independent samples 2-tailed t-test, $t(27) = .56$, $p = .58$) (Figure 5.7).

5.3.2 Transfer

Probability Transfer Test

In the probability transfer test the high- and low-probability statistical contingencies were exchanged. Improvement from the pre-training test to the untrained transfer test occurred (mixed two-way ANOVA, significant main effect of Training, $F(1,27) = 33.05$, $p < .001$), showing generalisation occurred from training to the probability transfer condition. Participants from the Different Contingencies group performed significantly higher on the probability transfer test compared to participants from the Main group (significant interaction effect of Training x Group, $F(1,27) = 5.02$, $p = .03$, and there was a significant main effect of Group ($F(1,27) = 5.76$, $p = .02$) indicating a significant difference between groups (Figure 5.8).

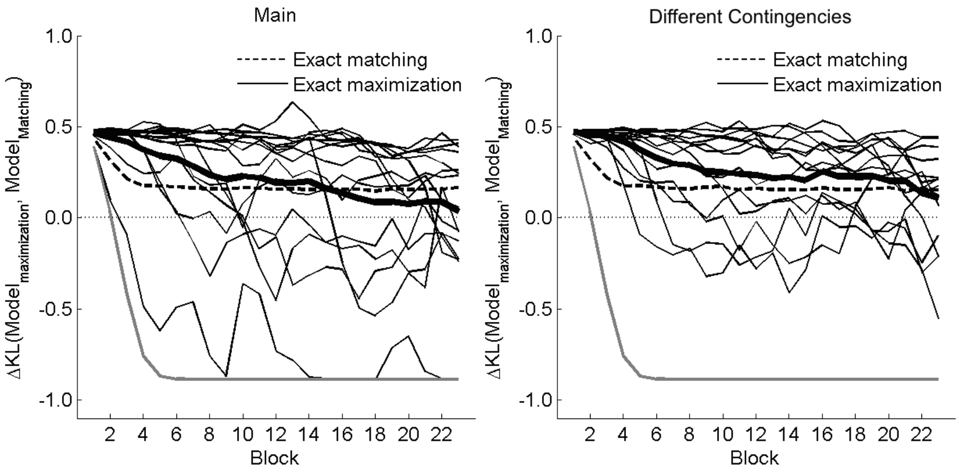


FIGURE 5.6: Strategy choice. Δ KL divergence between model matching and model maximization strategies for participants from the Main (n = 15) and Different Contingencies (n = 14) groups. KL divergence is a measure of how different the probability distributions that the subject is using to make predictions are from the probability distributions of predictions made using a perfect maximization strategy and a perfect matching strategy. The average curve is shown as a thicker line. Exact matching (dashed line) and maximization models (solid grey line) are plotted. More negative KL scores indicate a strategy closer to maximization.

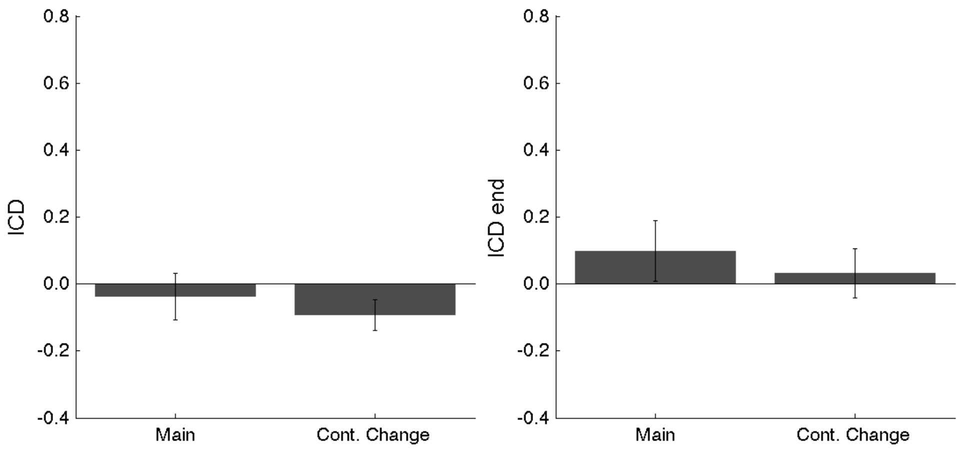
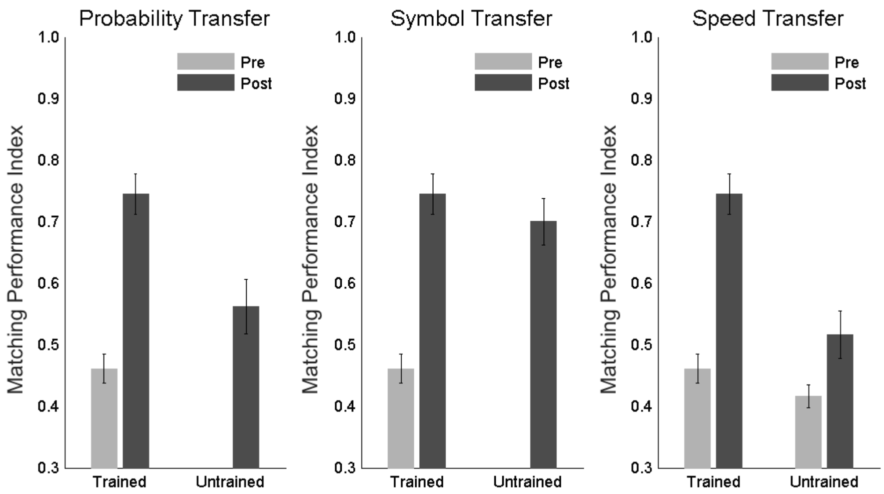
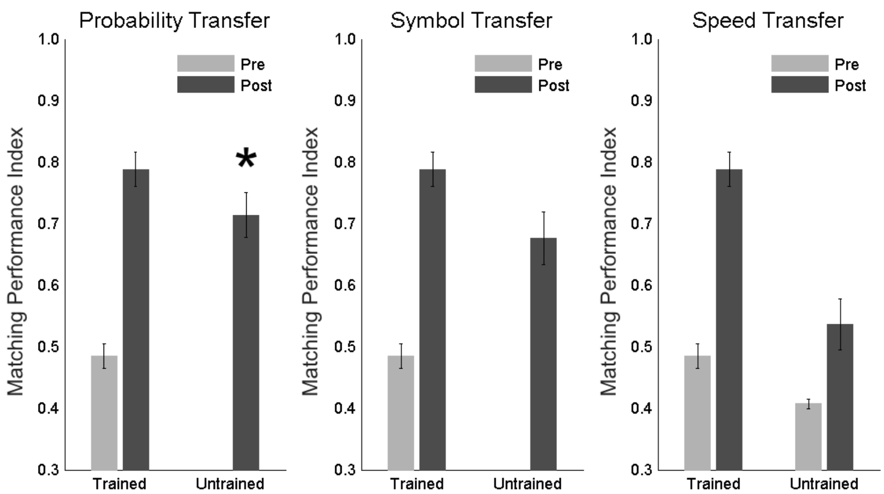


FIGURE 5.7: Mean strategy indices for participants from the Main (n = 15) and Different Contingencies (n = 14) groups. ICD and ICD end measure the signed area between the subjects' strategy curve and predictions made using a perfect matching strategy. Higher ICD and ICD end values indicate a strategy closer to maximization. ICD and ICD end values equal to zero indicate a perfect matching strategy.



(A) Main



(B) Different Contingencies

FIGURE 5.8: Transfer test performance for participants from the Main (n = 15) and Different Contingencies (n = 14) groups. Higher matching performance index values indicate more transfer.

Symbol Transfer Test

Improvement from the pre-training test to the untrained transfer test occurred (mixed two-way ANOVA, significant main effect of Training, $F(1,27) = 68.57$, $p < .001$), showing generalisation occurred from training to the symbol transfer condition. There was not an interaction effect (Training x Group, $F(1,27) = .82$, $p = .37$) nor a main effect of Group ($F(1,27) = .00$, $p = 1.00$), indicating that performance did not significantly differ between groups (Figure 5.8).

Speed Transfer Test

Improvement from the pre-training test to the untrained transfer test occurred (mixed two-way ANOVA, significant main effect of Training, $F(1,27) = 6.60$, $p = .02$), showing some generalisation occurred from training to the speed transfer condition. There was not an interaction effect (Training x Group, $F(1,27) = .01$, $p = .93$) nor a main effect of Group ($F(1,27) = .29$, $p = .60$), indicating that performance did not significantly differ between groups (Figure 5.8).

Direct Transfer Comparisons

This information is the same as presented in (Figure 5.8), but shows a direct comparison of the transfer conditions between groups.

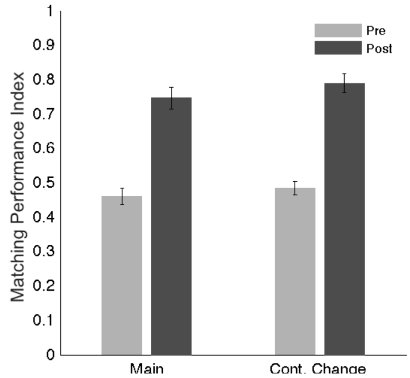
Mean performance in the probability transfer test was significantly different across groups (independent two-tailed t-test, $t(27) = -2.63$, $p = .01$); specifically, scores for the Different Contingencies group were greater than scores for the Main group (Figure 5.9(b)).

Mean performance in the symbol transfer test was not significantly different across groups (independent two-tailed t-test, $t(27) = .417$, $p = .68$) (Figure 5.9(c)).

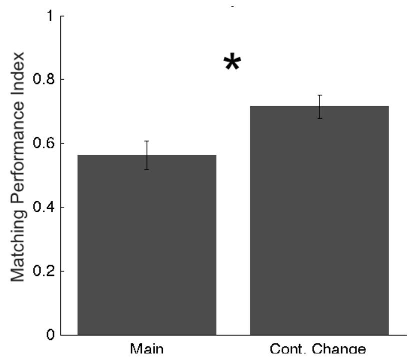
The speed transfer test did not show an interaction effect between Session x Group (mixed two-way ANOVA, $F(1, 27) = .28$, $p = .60$) nor a main effect of Group ($F(1,27) = .03$, $p = .87$), indicating that there were no significant differences between groups. However, there was a main effect of Session ($F(1,27) = 17.98$, $p < .001$) indicating some improvement on the transfer task from pre-training transfer test to post-training transfer test (Figure 5.9(d)).

5.4 Discussion

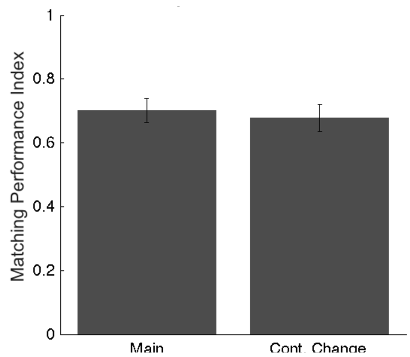
In this section, we asked whether people chose which learning strategies they employed based on the ease of execution. We specifically investigated two learning strategies, maximization and matching. To address this



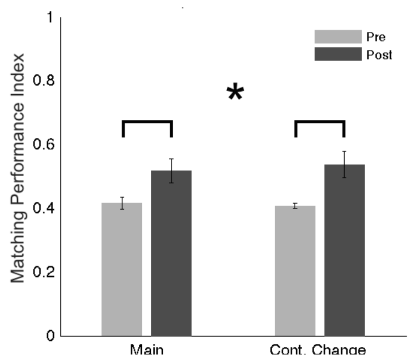
(A) Pre-training and post-training performance for comparison



(B) Probability Transfer



(C) Symbol Transfer



(D) Speed Transfer

FIGURE 5.9: Behavioural test and transfer test performance for participants from the Main (n = 15) and Different Contingencies (n = 14) groups. Higher matching performance index values indicate more transfer.

question, we compared two conditions that were roughly equated for maximizers but were of different difficulties for matchers. We hypothesized that in the condition where it was easier to match, more subjects would match, while the number of maximizers would be roughly similar across groups. This hypothesis led to the following set of predictions: the group where it was difficult to match ("Main" group) would have higher ICD and ICD end values, since they would have relatively fewer matchers (matchers have lower strategy indices than maximizers). Matching performance index scores were expected to be higher in the "Different Contingencies" group, since this group was expected to have more matchers. The Main group were expected to have roughly the same learning rate and transition time values as the Different Contingencies group since these values are similar across matchers and maximizers. Probability transfer scores from the Different Contingencies were expected to be higher because the low-probability statistical contingencies were simpler, because the transfer statistical contingencies were more similar to the training statistical contingencies in the Different Contingencies group, and because if there were more matchers some of them might have been making use of their low-probability knowledge from training. Scores on the other transfer tests were expected to be similar.

ICD and ICD end values and matching performance index scores were in the directions expected, but were not significantly different across groups. Learning rate and transition times were not significantly different across groups as expected. Probability, symbol, and speed transfer test performance was improved from pre-training test performance, demonstrating generalisation. As expected, participants from the Different Contingencies group performed significantly higher on the probability transfer test compared to participants from the Main group.

5.4.1 Matchers and maximizers

We asked the question of whether subjects choose their strategies based on the ease of execution. On the one hand, the results were in the expected directions, suggesting there may be merit to this theory. Adding support to this idea is the qualitative blurriness in defining "matchers" and "maximizers" in the strategy plots (Figure 5.6): subjects are seldom one or the other, but usually a combination of both. This is a function of the continuous nature of the ICD and ICD end measures, and shows that subjects rarely switch cleanly from being maximizers or matchers, but transition from one to the other as would be expected by a "ease of applicability" rationale.

On the other hand, the results were not close to significance: the differences between experimental groups was very small for the relevant measures (matching performance index and strategy indices). This could suggest that the choice of strategy, maximization or matching, is not a graded choice but absolute. Subjects will choose a strategy and do their best to stick with it. This would not be an unreasonable strategy if subjects were prone to being matchers or maximizers beforehand and were not aware of the difficulty of the low-probability statistical contingencies before they began. An immediate follow-up question to this conclusion is if subjects switch their strategies throughout their training, as they become more aware of contingencies. As discussed in Chapter 3, this information is captured in the strategy plots but is not captured in the summary statistics.

However, even in the strategy plots the value of ICD and ICD end do not fully capture the relationship between matching and maximization, because of the foreshortening of the area above the matching curve and the elongation of the area below the matching curve. This stretching makes it difficult to capture and quantify how these strategies interact. Moreover, the model used in this study cannot capture strategies that are not properly executed— if a subject is attempting a matching strategy but has only learned the probabilities within two contexts (e.g. the symbols that follow after "A" and "C") but has failed to grasp the probabilities for the remaining contexts (e.g. "B" and "D"), the subject's strategy choice line will appear more like that of a weak learner than a matcher. Future studies should develop measures that more completely capture the relationship between matchers, maximizers, and attempted matchers and maximizers, as well as capturing what it is about specific trials and sequences that may spark a subject's change in strategy.

5.4.2 Probability transfer test

One of the most interesting conditions in this set of experiments was the probability transfer test. Subjects in the Different Contingencies group were expected to outperform subjects from the Main group, which was what occurred. However, there were several different reasons why they were expected to perform well which are difficult to tease apart. First, the normal low-probability statistical contingencies from the Different Contingencies condition were intuitively easier to learn than those from the Main condition, so if subjects had learned nothing in training they should have performed better in the Different Contingencies probability transfer test. Second, the transfer probabilities from the Different Contingencies group were more similar to the original high-probability statistical contingencies that subjects were trained on in their respective groups, so if training was a factor, Different Contingency subjects would still be predicted to do better. Finally, matchers have to learn both high- and low-probability relationships, and it is possible that matchers consciously reverse those contingencies at test. Subjects who had learned the low-probability contingencies would be at an advantage over subjects who had just learned the high-probability contingencies, because the low-probability contingencies suddenly come to drive the sequence in the probability transfer task (when they switch to become the high-probability contingencies). If more subjects were in the Different Contingencies group were matchers, this could result in better performance in the probability transfer test. On the other hand, maximizers could also have this knowledge— as was argued in Chapter 3, subject responses can only capture how subjects act, not what they know.

All of these possible sources describing why subjects from the Different Contingencies group outperformed the Main group on this transfer test were confounded in this study. Future studies should develop tests that isolate these components. For example, to determine if training is irrelevant and it is only the difficulty of the transfer test that matters, subjects from both groups (Main and Different Contingencies) should be tested on the same transfer test and should perform equivalently. To determine if it is only the similarity of the training and transfer statistical probabilities that matters, subjects in both groups should be trained on the Main transfer test statistical probabilities, and tested on their current transfer tests with different symbols: subjects in Main condition should outperform the Different Contingencies subjects since the oddly-configured loop will be more familiar than the more straightforward loop. It is harder to design a test to determine if matchers use their low-probability knowledge to perform the probability transfer test. One interesting analysis would be to compare performance from matchers and maximizers within the same group, being careful to ensure that matchers and maximizers can be sorted out neatly and not be either confused with each other or combined with subjects who have a poorer grasp of some of the statistical contingencies. This analysis, however, only works assuming that maximizers have no low-probability knowledge, so a better test might be to first test subjects' knowledge of the low-probability contingencies and then separately examine their probability transfer test generalisation results.

5.4.3 Conclusions

In this section we asked how subjects used two important strategies for learning: matching and maximization. Specifically, we hypothesized that subjects would differentially employ these strategies depending on ease of use. To test this, we compared two experiments of similar difficulty for maximizers but of different difficulty for matchers. We expected that in the group where matching was easier more subjects would be matchers. We observed trends in this direction but far from significant differences between groups. This suggests the need for further work exploring which factors do govern how these two strategies are employed, as ease of application does not necessarily seem to be the metric that subjects use.

Future work could include probing whether subjects have strong tendencies toward maximization or matching before the experiment begins and how flexible subjects are at changing their strategies. Moreover, stronger strategy measures should be developed that allow better characterisation of how maximization and matching interact within hybrid strategies, and also how to model strategies from subjects with incomplete knowledge. Additionally, the finding that subjects from the Different Contingencies group were able to outperform the Main group on the probability transfer test gives rise to many more follow-up studies investigating about how subjects learn and why this result should occur.

Comparison of these experimental groups revealed that how subjects choose strategies is not as simple as choosing the strategy that is easiest, which perhaps could have been predicted based on the fact that not all subjects are maximizers. The fact that not all subjects are maximizers also suggests that subjects do not choose strategies based on accuracy-based optimality. Further investigation into the details of maximization and matching will continue to seek the factors that drive how subjects choose to learn and act.

Chapter 6

General Discussion

Statistical learning, the ability to learn the statistical relationships between items and events, is one of our most powerful mechanisms for navigating a complicated world. From learning language to predicting what comes next, this is an essential ability spanning modalities and domains. We sought to investigate how subjects direct their statistical learning based on situational demands: whether subjects would engage different strategies based on the task structure. To investigate this question we introduced three manipulations.

We first exposed participants to feedback, expecting that the acknowledgement of monitoring coupled with the explicit introduction of desired outcomes would induce subjects to change their strategies. Subjects did indeed change their strategies, maximizing more when feedback that encouraged maximization (both accuracy-based and maximization-based) was applied. Interestingly, subjects receiving intermittent feedback encouraging matching did not seem to maximize or match more than subjects not receiving any feedback, a result that encourages future study about the role of directed feedback.

We then introduced disruptions in the fundamental property of temporal relations. We expected subjects to perform more poorly and develop poorly-performing strategies; but instead, subjects in the groups with the most jitter adopted maximization strategies, perhaps due to a "hardness hypothesis" in which subjects became overwhelmed by difficult stimuli and chose to optimize their strategy from the beginning.

Finally, we introduced a sequence task with a new structure, to determine if subjects would change their strategies when the relative ease of certain strategies was different between groups. Subjects did not change their strategies in this case but did have superior performance on a transfer test, showing that they understood the ease of the low-probability contingencies but did not apply this knowledge to their strategy use. This result is curious, as subjects do not appear to base their strategy choices on ease of use or accuracy-based optimality (otherwise everyone would be a maximizer), which raises further questions into the factors that drive peoples' approaches to learning and prediction.

We examined the strategies that subjects used to direct their statistical learning, but what learning occurred in the first place? In the following sections, we examine what rules and content subjects were acquiring such that their specific strategies emerged.

6.1 Abstraction: Symbol Transfer Test

When people learn a complex sequence like the one in this study, to what degree of abstraction do they acquire rules? In fact, subjects across all groups significantly improved from the pre-training test to the symbol transfer test (3.7, 4.7, 5.8). This demonstrates that no matter what the task demanded, subjects in this task learned rules that generalised over the surface feature of shape.

This finding of abstraction over shape is not unprecedented; in visual statistical learning studies, Turk-Browne, Jungé, and Scholl (2005) showed that subjects could classify triplets of symbols and generalise over a surface feature like colour, while Turk-Browne, Isola, et al. (2008) observed that subjects could classify triplets based on colour and generalise over shape. Subjects can transfer to new letter sets in artificial grammar learning and artificial language learning paradigms, showing abstraction over shape (e.g. Reber (1967), Marcus, Vijayan, et al. (1999), and Gerken (2010)).

However, the finding of generalisation observed in this study was also not guaranteed. The current study can be compared with Gómez (1997), one of the few studies that uses a similar transfer test. In one of their experiments, Gómez (1997) used a sequential artificial grammar learning paradigm in which the untrained letters appeared on the screen one at a time in the transfer test. This transfer test is analogous to the symbol transfer test used in this study. Gómez (1997) also had an experiment using the standard artificial grammar learning paradigm in which strings containing untrained letters were presented simultaneously on the screen in the transfer test. Interestingly, subjects could transfer their knowledge in the standard artificial grammar learning transfer test but not in the sequential test. The fact that generalisation over shape was not observed in Gómez (1997) is particularly salient given the task complexity of a sequential artificial grammar transfer test is similar to that of the current study. It is possible that subjects in the Gómez (1997) study may have been able to transfer their knowledge with extensive training, as our subjects underwent multiple training sessions with sleep consolidation. In the current study, the success of transfer in the symbols transfer test across all sets of experimental groups is strong support that subjects' statistical learning was abstract over the feature of shape. Future studies should investigate which surface features will be persistently generalised over across paradigms, and what components of different statistical learning tasks drive the specificity of learning.

6.2 Knowledge vs. Decision-Making

This question of what information subjects have acquired is distinct from what information subjects choose to behaviourally demonstrate. One of the drawbacks of this study is that while matchers show their learning of both low- and high-probability statistical contingencies, maximizers by definition of their strategy choice only demonstrate their high-probability knowledge. Future studies must employ precise response measures that are calibrated to target specific questions; these probes need not be explicit, but they must test that subjects can predict items using low-probability distributions as well as high-probability distributions. From another angle, something that would have helped in linking decision-making strategies with learning would have been to analyse the results from the transfer tests for matchers and maximizers separately. Unfortunately, another critique of this study is that the border between matchers and maximizers is continuous and condensed, which makes it difficult to separate matchers and maximizers. However in the future when the separation would be more clear-cut, we could observe whether matchers and maximizers have different abilities to generalise their knowledge, which would be an indication of what information they had learned during training. Future work will require carefully designed questions and models of the information subjects are representing. We are always limited by our input analysis measure—behaviour— but careful controls can reduce the gap between subjects' responses and our hypotheses of what subjects are representing.

6.3 Future Studies

The question of what strategies people use to solve probabilistic problems is an interesting one. From the results of this study, several follow-up questions arise that would help reveal the role of these strategies. First, do people match before they maximize (as would be supported by the idea that matchers learn more information that maximizers)? Do subjects switch between maximization and matching strategies, or do they adopt characteristics of both? Do subjects pursue one strategy for a given amount of time before they switch, or is a switch externally driven? Which specific sequences may bias a subject towards one strategy or another? Turning the question on its head: does the decision to adopt a strategy influence what information a subject acquires, or does information-acquisition drive the strategy?

There are also a number of questions one could ask with regards to individual differences in adopting strategies. Are some strategies followed first? Are some people born matchers? Do behavioural characteristics predict which strategies a subject uses regardless of task demands? Correlation studies with behavioural characteristics like IQ or mathematical training would be useful in addressing this last question. Another approach would be to correlate which strategies subjects used in this task to what strategies subjects would use on a binary prediction task (e.g. the marbles task), a simpler task where maximizing and matching strategies are also employed. Additionally, it would be interesting to compare subject performance on multiple statistical learning tasks to see if certain characteristics across tasks might predict performance on this one.

This paradigm offers intriguing insights into strategy use in a complex paradigm mimicking how subjects engage statistical learning in real life. Long-term, strategic decision-making is an important component in understanding how people interact with a probabilistic world, but the literature in statistical learning often does not operate within this framework, and the literature in maximizing and matching often uses simple binary prediction paradigms. This study integrates both domains and raises interesting questions but has methodological limitations; future studies should continue to study the interaction between these mediums. These studies could continue to investigate task demands like the ones implemented in this study (feedback, temporal jitter, structural contingencies) but also others used in statistical learning paradigms (e.g. colour, interleaved sequences, occasional deterministic sequences). Pursuing these questions would help describe which particular components of this paradigm drove the unpredicted results by drawing comparisons between statistical learning experiments.

6.4 Conclusions

This study investigated the strategies involved in statistical learning. Specifically, we probed whether subjects changed how they made their predictions— using a maximization or matching strategy— depending on the task structure. We observed that subjects often did change their strategies (sometimes in unpredicted directions) and many follow-up questions can be raised. Future experiments can provide further detail into what features motivate subjects to choose strategies in probabilistic settings, addressing the deeper question of how people effectively harness their ability to learn complex, statistical information.

Appendix A

Individual Differences Analyses

In this study we were interested in what strategies subjects adopted, and an important factor in strategy choice could be individual differences. We thus analysed relationships among individual subjects' strategy scores and performance scores, learning rates and independent cognitive test results to discover if there were factors that could predict strategy at an individual level as opposed to a group level.

A.1 Learning Profile and Strategy Profile Correlations

In previous work (Wang et al., in review), learning rate and strategy indices were positively correlated, indicating that subjects who learned quickly tended to use strategies closer to maximization. Specifically, learning rates (higher values indicate faster learners) and ICD / ICD end (strategy indices, higher values indicate more maximization behaviour) were correlated. This relationship between learning rate and ICD / ICD end was maintained in the current set of experimental groups (ICD vs. LR, $r(98) = .51$, $p = 4.7e{-8}$; ICD end vs. LR, $r(98) = .47$, $p = 9.4e{-7}$, Figure A.1). However, it is interesting to note that some of this correlation appears to be driven by subjects with very low learning rates and negative ICD and ICD end scores—these are subjects who did not fully learn the task and did not adopt either a maximization or matching strategy. Therefore, one should be cautious about making conclusions concerning whether it is the maximization strategy specifically that correlates with learning rate as opposed to maximization and matching strategies that are correlated with fast learning.

Similarly, performance on the post-training test (measured in matching performance index) correlated positively with learning rate and strategy indices: Post-training test vs. LR: $r(98) = .35$, $p = .00042$; Post-training test vs. ICD: $r(98) = .52$, $p < 3.2e{-8}$; Post-training test vs. ICD end: $r(98) = .56$, $p =$

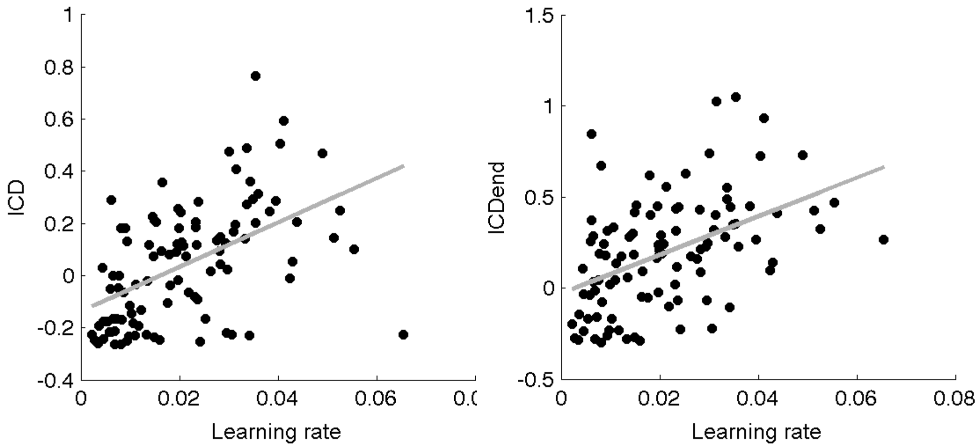


FIGURE A.1: Correlations between learning rate and strategy indices. Data from 100 participants was pooled across experimental groups: Main (n = 15), Different Contingencies (n = 14), Jitter (n = 14), Augmented Jitter (n = 15), Feedback (n = 17), Maximization Feedback (n = 14), and No Feedback (n = 11). Learning rate is the slope of the learning sigmoid curve; higher learning rates indicate faster learning. ICD (integral curve difference) and ICD end measure the signed area between the subjects' strategy curve and predictions made using a perfect matching strategy. Higher ICD and ICD end values indicate a strategy closer to maximization. ICD and ICD end values equal to zero indicate a perfect matching strategy.

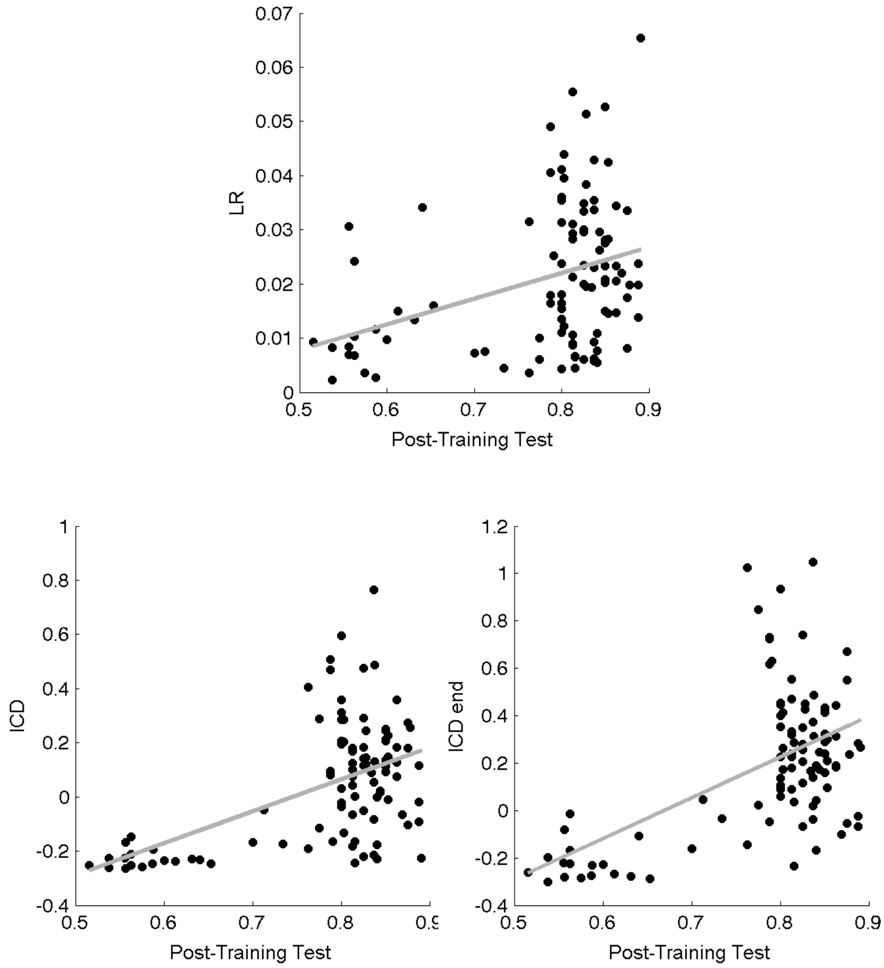


FIGURE A.2: Post-training test performance, measured in matching performance index, correlations with learning rate and strategy indices. Data from 100 participants was pooled across experiments: Main (n = 15), Different Contingencies (n = 14), Jitter (n = 14), Augmented Jitter (n = 15), Feedback (n = 17), Maximization Feedback (n = 14), and No Feedback (n = 11). No weak learners were included as usual; data suggest groupings within included subjects. Post-training scores are matching performance index, where higher values indicates more matching behaviour. Learning rate is the slope of the learning sigmoid curve; higher learning rates indicate faster learning. ICD (integral curve difference) and ICD end measure the signed area between the subjects' strategy curve and predictions made using a perfect matching strategy. Higher ICD and ICD end values indicate a strategy closer to maximization. ICD and ICD end values equal to zero indicate a perfect matching strategy.

2.2e-9 (Figure A.2). Here, learning rate and strategy indices correlate with a measure of matching behaviour, further enforcing the idea that adoption of either a maximizing or matching strategy correlates with fast learning. Further, the correlations again appear to be driven by a smaller subset of subjects with low post-training scores, low learning rates, and low ICD / ICD end scores. When subjects are not performing the task effectively, indicated by low post-training scores and low learning rates, the model has difficulty capturing whether subjects are following a strategy closer to matching or maximization (where lack of modellable strategy is indicated by very low ICD and ICD end scores). Future studies might try to better capture subjects' behaviour when they are not using either a maximizing or matching strategy effectively, but may well still be using a strategy closer to one or the other.

A.2 Long-Term Learning Maintenance

Would the strategies that subjects employed be maintained beyond the short period of the learning study? Participants in statistical learning studies are rarely recalled to see if their learning is maintained in the long term. When subjects are recalled, results after 24 hours are considered a long-term effect (Kim et al., 2009). However, the current study was an intensive five-day paradigm, therefore learning was expected to be maintained over a longer period. Like in Baker et al. (2014), subjects were recalled after more than a month on average to determine if statistical learning had been maintained.

Strategy use was not directly examined, but performance on the post-training test measured in matching performance index was analysed. If matching performance index scores were maintained across subjects over time, this would be an indication that subjects maintained their matching and maximization strategies beyond the learning period. In fact, performance on the post-training test was maintained over time. 33 participants from the Main (n = 11), Different Contingencies (n = 9), Jitter (n = 8), and Augmented Jitter (n = 5) experimental groups completed another post-training test (mean time elapsed since previous post-training test: 50.3 ± 6.0 days). Means (post- vs. recalled post-training test) were not significantly different across Session (paired-samples t-test, $t(32) = 1.61$, $p = .12$). Mean performance on post-training test = $.79 \pm .02$; performance on recalled post-training test = $.78 \pm .02$.

Moreover, we asked whether strategy (as measured indirectly through performance matching index) would be maintained over time on the transfer test that most subjects successfully completed: the symbol transfer test. Maintenance of scores on the symbol transfer test would indicate that subjects continued to use the same matching and maximization strategies for

Cognitive Tests	ICD	ICD end
VSTM	-.01 (.91)	-.0028 (.98)
UFOV	-.28 (.0049)	-.20 (.048)
Episodic	.28 (.055)	.29 (.050)
ProbRev (persev)	-.0041 (.98)	-.0060 (.97)
ProbRev (pswitch)	.14 (.35)	-.16 (.29)

TABLE A.1: Correlations for cognitive tests and strategy indices. Data was pooled across experimental groups. For the Visual Short-Term Memory ("VSTM") task and Useful Field of View ("UFOV") task, data was pooled across 100 participants: Main (n = 15), Feedback (n = 17), Maximization Feedback (n = 14), and No Feedback (n = 11), Jitter (n = 14), Augmented Jitter (n = 15), Different Contingencies (n = 14). For the Episodic Memory ("Episodic") and Probabilistic Reversal ("ProbRev") tasks, data was pooled across 46 participants: Main (n = 12), Maximization Feedback (n = 2), Jitter (n = 8), Augmented Jitter (n = 15), Different Contingencies (n = 9). R-values for correlations are included with p-values in parentheses. ICD (integral curve difference) and ICD end measure the signed area between the subjects' strategy curve and predictions made using a perfect matching strategy. Higher ICD and ICD end values indicate a strategy closer to maximization. ICD and ICD end values equal to zero indicate a perfect matching strategy. Higher VSTM scores indicate better short-term memory, lower UFOV scores indicate better selective attention, higher episodic memory scores indicate better episodic memory, and lower probabilistic reversal scores indicate better performance. Specifically, lower perseverance scores ("persev") indicates subjects switched rules quickly when appropriate (switched rules when task structure changed), and lower probability switch scores ("pswitch") indicate that subjects persisted in pursuing rules when appropriate (persisted despite intermittent negative feedback).

transfer over time. The result was that performance on the symbol transfer test (measured in matching performance index) was maintained over time for the small set of subjects who were retested on the transfer test. Participants in the Augmented Jitter (n = 5) group completed another post-training test (mean time elapsed since previous post-symbol test: 57.0 ± 5.7 days). Means (post- vs. recalled post-training symbol test) were not significantly different across Session (paired-samples t-test, $t(4) = -.68$, $p = .54$). Mean performance on symbol transfer test = $.79 \pm .02$; performance on recalled symbol transfer test = $.81 \pm .01$.

A.3 Cognitive Test Correlations

We sought in this experiment to understand why subjects used maximization or matching strategies. Analysing correlations between strategy use and independent cognitive tests provides insight into the factors that govern strategy choice. We compared the results from subjects' strategy measures with four cognitive tests: Visual Short-Term Memory (VSTM), Useful Field of View (UFOV), Episodic Memory, and Probabilistic Reversal.

A.3.1 Methods

Two strategy indices, integral curve difference (ICD) and ICD end (ICD calculated over the last two blocks of training rather than across all training blocks) were used.

Four cognitive tests were used: a VSTM task to probe short-term memory, a UFOV task to probe selective attention, an episodic memory task, and a probabilistic reversal task. (Note that the RSVP task was reserved for screening due to ceiling effects.) The VSTM and UFOV tasks were described in Chapter 2; in the VSTM task, subjects saw a number of coloured dots and had to recall the colours at test, while in the UFOV task subjects saw a target image in the center of their vision and another at their periphery, and at test had to recall the identity of the target image at the center and the location of the image at the periphery. The episodic memory and probabilistic reversal tasks are described below.

Strategy indices were correlated with performance on all four cognitive tests.

Episodic Memory Task

A treasure-hunt task was used to test episodic memory (Cheke, Simons, and Clayton, 2016). Participants were presented with scenes and instructed to move and "hide" several food item images using the landmarks in each scene. In each trial, two different scenes were used; four items were presented within each scene. Each item was presented twice within a given scene across two consecutive hiding periods named "day 1" and "day 2". Participants were instructed to not hide any two items in the same place across days. In a recall period, subjects were required to indicate where they had hidden each item in each scene during each of the "days". The number of items whose positions for each day were successfully retrieved was recorded. Participants completed a training session with feedback before completing two test trials without feedback. 46 participants completed the Episodic Memory tasks.

Probabilistic Reversal Task

The probabilistic reversal task was used to assess cognitive flexibility (Murphy et al., 2003). Participants were presented with two images of stripes— one blue, one yellow— in each trial. Subjects clicked one of the coloured images with a mouse and were given trial-by-trial feedback (correct/incorrect) to encourage them to continue choosing the image of specified colour. Feedback was misleading 25% of the time. Reversal of the stimulus-reward contingency took place after 40 trials (the colour that was previously "incorrect" became "correct" and vice-versa). Subjects were instructed to "choose the image that is most often correct and this rule may change." Two measures of cognitive flexibility were recorded: "perseverance" marked how many trials the subject continued choosing the colour that had been correct before the contingency-reversal after the reversal (and accompanying feedback changes) occurred, and "probability of switching" described the number of times the subject chose the incorrect colour immediately in response to negative feedback (when they should have adhered strictly to a colour rule). 46 participants completed the probabilistic reversal task.

A.3.2 Predictions

All cognitive tests— VSTM, UFOV, episodic memory, and probabilistic reversal (measures: perseverance and probability of switching)— were correlated with strategy measures (ICD and ICD end).

It was predicted that high performance on the memory and attention tasks would be correlated with higher ICD and ICD end scores, indicating that increased memory and attentional abilities would facilitate subjects adopting maximization and matching strategies. It was hypothesized that memory would have a stronger impact on strategy choice (matchers in particular need to have a large memory capacity to recall probability distributions) than attention, so the correlations between ICD and ICD end and the memory tests would be stronger than the correlations between ICD and ICD end and the attention task.

Specifically, ICD (integral curve difference) and ICD end measure the signed area between the subjects' strategy curve and predictions made using a perfect matching strategy. Higher ICD and ICD end values indicate a strategy closer to maximization. ICD and ICD end values equal to zero indicate a perfect matching strategy. The memory tasks were VSTM and episodic memory. Higher VSTM scores indicate better short-term memory and higher episodic memory scores indicate better episodic memory. Therefore ICD / ICD end and VSTM were expected to be positively correlated and ICD / ICD end and episodic memory were expected to be positively corrected. The attention task was UFOV. Lower UFOV scores indicate better selective attention. Thus, ICD / ICD end and UFOV scores were expected to be negatively correlated, but to a lesser degree than the memory tasks.

For the probabilistic reversal task, it was predicted that higher performance would correlate with more maximization and matching strategies. Lower probabilistic reversal scores indicate better performance. Specifically, lower perseverance scores indicate that subjects switch rules quickly when appropriate (change strategies when faced with changing structure), and lower probability switch scores indicate that subjects persist in pursuing given rules when appropriate (maintain strategies in the face of intermittent negative feedback). If subjects tend to change rules to match the task structure they observe (loosely what the perseverance score measures for the probabilistic reversal task), they should adopt maximization and matching strategies appropriately, resulting in high ICD and ICD end scores. Thus, perseverance scores and ICD / ICD end should be negatively correlated. Also, if subjects persist in pursuing rules despite intermittent negative feedback (loosely what the probability switch score measures for the probabilistic reversal task), subjects should correctly persist in their matching and maximization strategies even when they occasionally fail to predict the next symbol correctly. This indicates that probability switch scores and ICD / ICD end scores should be negatively correlated.

If the predicted correlations for all cognitive tests are strong, this would be an indication that strategy choice is as much a measure of individual differences as group-level manipulations. Strategy choice could be predicted on the basis of independent cognitive abilities.

A.3.3 Results and Discussion

Two scores correlated with p-values close to .05: UFOV scores and episodic memory scores. UFOV scores were negatively correlated with ICD and ICD end (Table A.1). Lower UFOV scores indicate better performance, so this correlation was in the direction expected: subjects with strong selective attention abilities tended to be maximizers and matchers. UFOV is an attentional task, so attention thus played a more important role in strategy choice than predicted. Episodic memory scores were positively correlated with ICD and ICD end. Higher episodic memory scores indicate higher performance, so this correlation was also in the direction expected: subjects with strong episodic memory abilities tended to be maximizers and matchers. If maximizers and matchers could be neatly separated in this paradigm, it would have been interesting to analyse the correlation between ICD and ICD end and episodic memory scores for maximizers and matchers independently since higher episodic memory ability might have been expected for the matchers (who had to memorize entire probability distributions; this might be the case for maximizers as well but was necessary for matchers).

It is intriguing that scores from the VSTM task, which is also a memory task, did not correlate with ICD and ICD end. It speaks to an interesting difference between episodic memory and short-term memory. In the sequential learning task, one would expect both abilities were important— the ability to recall specific episodes of sequences and also the ability to recall the previously-presented symbol. The finding that episodic memory scores correlate with strategy but VSTM scores do not may suggest that remembering past episodes of symbols— whatever length of time that entails in the sequential learning task— may be more predictive than the ability to remember the just previous symbol, which is essential but perhaps not as rare an ability. This finding is difficult to interpret, however, because analogies between episodic memory as defined in the episodic memory task and episodic memory as defined in the sequential learning task are tentative; the same is true for short-term attention analogies. One might have predicted that all cognitive tests would correlate with ICD and ICD end, implicating an underlying variable driving high performance in general like IQ. However, the lack of correlation with VSTM scores proposes this not the case, suggesting that there is merit in analysing individual cognitive abilities for correlations with strategy use.

Finally, it was unexpected that probabilistic reversal task scores did not correlate with strategy indices. This task has previously revealed behavioural differences between depressed patients and controls in Murphy et al. (2003), so it was expected that differences would be observed here. However, there were no patient groups in this study, so perhaps lack of sensitivity to the proposed cognitive abilities due to low individual variability could have led to the lack of correlations. Future studies should continue to investigate the link between the probabilistic reversal task and strategy use, because conceptually the abilities investigated in the probabilistic reversal task— tendencies to switch strategy with regards to task structure, and tendencies to persist in response to negative feedback— seem incredibly relevant to strategy decisions. Additional cognitive tasks should be developed aimed at isolating these abilities just in case they cannot be captured in non-clinical subjects participating in the probabilistic reversal task used here.

The independent cognitive tests produced an interesting set of results suggesting that individual capabilities in specific tasks may be able to predict strategy use. Future work should continue investigating which abilities may be predictive of strategy behaviour, perhaps controlling for a variable like IQ as well. Future work should also consider the relative impact of individual variables compared to group-level manipulations— in this study, subject numbers were often too small to find correlations when subjects were divided into experimental groups, but this interaction would be intriguing in seeking to understand how flexible subject strategies are to large-scale interventions like the ones proposed in the current study.

Appendix B

Supplementary Methods

The following section contains the supplementary methods from a previous study submitted for publication in Wang et al. (in review). The original submission has been lightly edited to reflect the methods used in the current study.

Sequence design

To generate probabilistic sequences with different complexity levels, we used a temporal Markov model and systematically varied the memory length (i.e. context length) of the sequence. The model consists of a series of symbols, where the symbol at time i is determined probabilistically by the previous k symbols (context). We refer to the symbol presented at time i, $s(i)$, as the target and to the preceding k-tuple of symbols $\big(s(i-1), s(i-2), ..., s(i-k)\big)$ as the context. The value of k is the order, or level, of the sequence: $P\big(s(i)|s(i-1), s(i-2), ..., s(1)\big) = P\big(s(i)|s(i-1), s(i-2), ..., s(i-k)\big), k < i$.

- The simplest $k = 0$th order model is a random memory-less source. At each time step i, this generates a symbol according to symbol probability $P(s)$, without taking the context (i.e. previously generated symbols) into account.
- The $k = 1$ model generates symbol $s(i)$ at each time i conditional on the previous symbol $s(i-1)$. This introduces a memory in the sequence, i.e., the probability of a particular symbol at time i strongly depends on the preceding symbol $s(i-1)$. Unconditional symbol probabilities $P\big(s(i)\big)$ for the case $k = 0$ are now replaced with conditional ones, $P\big(s(i)|s(i-1)\big)$.

At each time point, the symbol that follows a given context is determined probabilistically, thus generating stochastic Markov sequences. The underlying Markov model can be represented through the context-conditional target probabilities. We used 4 symbols that we refer to as items A, B, C and D. The correspondence between items and symbols was counterbalanced across participants. For level-0 sequences (random sequences) each symbol had an equal probability of occurrence. For 1st-order sequences, the target depended on the preceding item. Given a context (the last seen symbol) only one of two targets could follow; one had a high probability and the other low probability (e.g., 80% vs. 20%). For example, when Symbol A was presented, only symbols B or C were allowed to follow, and B had a higher probability of occurrence than C.

All sequences were pre-generated. To ensure that sequences in each block were representative of the Markov model per sequence type (training, test, random (presented within the test segments of test block / random block / test block), probability transfer, symbol transfer, and speed transfer), we generated 200 Markov sequences per sequence type comprising 672 items per sequence. We then estimated the Kullback- Leibler divergence (KL divergence)

$$\mathrm{KL} = \sum_{c} Q(c) \log \frac{Q(c)}{P(c)}$$

for each presented sequence $P(c)$ compared to the ideal Markov model $Q(c)$ across the probabilities of all the conditions c (individual targets at level-0 or context-target contingencies at level-1) and selected the number of blocks plus one sequences with the lowest KL divergence (i.e. these sequences matched closely the Markov model per sequence type and level). This process was repeated for each session. The sequences presented to the participants during the experiments were selected randomly from these sequence sets. (Note: when calculating KL-divergences, we replaced zero/negligible probabilities with a small value 0.001 and then renormalized.)

Data analysis

Matching performance index: We assessed participant responses in a probabilistic manner. For each context, we computed the absolute distance between the distribution of participant responses $P_{resp}(s|context)$ and the distribution of presented targets $P_{pres}(s|context)$ estimated across 60 trials per block:

$$AD(context) = \sum_{t} |P_{resp}(s_t|context_t) - P_{pres}(s_t|context_t)|,$$

where t is trial index and the target s is from the symbol set A, B, C and D. We quantified the overlap between these two distributions by computing a matching performance index per context:

$$PI(context) = \sum_t \min\big(P_{resp}(s_t|context_t), P_{pres}(s_t|context_t)\big).$$

Note that $PI(context) = 1 - AD(context)/2$. The overall matching performance index is then computed as the average of the matching performance indices across contexts, $PI(context)$, weighted by the corresponding context probabilities:

$$PI = \sum_{context} PI(context) \cdot P(context).$$

To compare across different levels, we defined a normalized matching PI measure that quantifies relative participant performance above random guessing. We computed a random guess baseline; i.e. matching performance index PI_{rand} that reflects participant responses to targets with equal probability for each target for a given context for level-1 (matching $PI_{rand} = 0.45$). To correct for differences in random-guess baselines across levels, we subtracted the random guess baseline from the matching performance index, (matching $PI_{normalized} = PI - PI_{rand}$).

Performance accuracy: To compare the performance for both structured and random sequences, we calculated accuracy (percent correct) across trials; that is, we computed whether the test stimulus was correctly predicted or not— i.e. the participants response matched the exact outcome based on the pre-defined sequences. This value, however, was not included in the figures.

Model description

We developed a computational model that tracks human predictions as they evolve over time. Using this model, we extract changes in performance over time that relate to: (1) learning the memory order that governs the sequences (i.e., identifying the context length); and (2) learning to generate a prediction about the next item given the current context. We refer to these two components as context-length and predictive contingency model, respectively.

First, we describe the context-length model, which considers participants' responses as a weighted combination of multiple Markov processes. This modelling approach enables us to track participants' learning in a unified framework. When participants are first exposed to the sequences, we reason that their responses will be driven by random guesses, which corresponds to a special setting of the zero-order (memory-less) Markov model.

We reason that participants' responses will be later refined after having observed that some symbols are presented more frequently than others in a given context. At trial t, our model tracks the context-length learning using two components corresponding to no memory (zero-order Markov model $\mathcal{M}_t^0$ using empty context $\emptyset$ of zero length) or shallow memory (first-order Markov model $\mathcal{M}_t^1$ using context $\mathcal{C}_t^1$ of length one):

$$\mathcal{M}_t = \sum_{k\in\{0,1\}} w_t^k \cdot M_t^k, \quad \text{subject to} \quad \sum_k w_t^k = 1 \quad \text{with} \quad w_t^k \geq 0, \tag{1}$$

where w_t^k is the mixture coefficient at trial t of the component model $\mathcal{M}_t^k$ and represents the probability of using $\mathcal{M}_t^k$ for prediction at trial t. The mixture coefficients w_t^k can be thought of as expressing the strength at trial t of individual components $\mathcal{M}_t^k$ in the overall mixture model. By tracking the mixture coefficients over trials, we dynamically assess whether participants' responses can be accounted for by the different Markov structures.

Given the history of last seen symbols, model (1) gives the following probability to a target symbol s:

$$p_t(s|\{\emptyset, \mathcal{C}_t^1\}) = w_t^0 \cdot p_t^0(s|\emptyset) + w_t^1 \cdot p_t^1(s|\mathcal{C}_t^1). \tag{2}$$

The predictive contingency distribution allows us to quantify how the participants' responses compare to the context-conditional probabilities of the generating Markov model.

To track a participant's responses over time with model (1), mixture coefficients as well as the mixture components themselves are updated after each response. The model calculates whether the participant's response is more likely to be driven by a particular mixture component (e.g. Markov model of order one), and updates the weight for the components accordingly. The model updating is implemented in Bayesian terms: given the participant's response at trial t, each mixture coefficient w_{t-1}^k is updated proportionally to both the current coefficient value ('strength' of $\mathcal{M}_{t-1}^k$) and the likelihood of $\mathcal{M}_{t-1}^k$ given the participant's response. Considering w_{t-1}^k and w_t^k as the prior and posterior for the model order k, Bayesian update of w_t^k reads:

$$W_t^k = \frac{w_{t-1}^k \cdot p_{t-1}^k(\hat{s}_t|\mathcal{C}_t^k)}{\sum_k w_{t-1}^k \cdot p_{t-1}^k(\hat{s}_t|\mathcal{C}_t^k)}. \tag{3}$$

Similarly, by applying the Bayes rule, the context-conditional predictive distribution is obtained:

$$P_t^k(s|\mathcal{C}_t^k) = \frac{\delta_{s,\hat{s}_t} \cdot p_{t-1}^k(s|\mathcal{C}_t^k)}{\sum_s \delta_{s,\hat{s}_t} \cdot p_{t-1}^k(s|\mathcal{C}_t^k)}. \tag{4}$$

Here we introduce a noise model δ representing uncertainty in the participant's prediction. If we assumed that the participant is certain about their prediction, δ would be a noise-less model (delta-function),

$$\delta_{s,\hat{s}_t} = \begin{cases} 1 & \text{if } s = \hat{s}_t, \\ 0 & \text{if } s \neq \hat{s}_t. \end{cases}$$

As the participants learn during the sequence presentation, we assume that the tracking model (1) of the participant's responses changes smoothly over time. This is represented by a partial adaptation of the model parameters towards the ideal Bayesian updates (3) and (4):

$$w_t^k = w_{t-1}^k + \eta_w \cdot (W_t^k - w_{t-1}^k) \tag{5}$$

and

$$p_t^k(s|\mathcal{C}_t^k) = p_{t-1}^k(s|\mathcal{C}_t^k) + \eta_p \cdot w_t^k \cdot \left(P_t^k(s|\mathcal{C}_t^k) - p_{t-1}^k(s|\mathcal{C}_t^k)\right), \tag{6}$$

where $0 < \eta_p < 1$ and $0 < \eta_w < 1$. Here we introduced tuning parameters η_p, η_w (i.e. the adaptation rate) to calibrate our tracking model and ensure smooth learning curves. The case of $\eta_p = \eta_w = 1$ corresponds to ideal Bayesian updating. From the learning dynamics point of view, for $\eta_p, \eta_w < 1$, the change of $\mathcal{M}_t$ over t is effectively smoothed. In our study, we assume the adaptation rates attain the same value, $\eta_p = \eta_w = \eta$, and set η to 0.04. The qualitative criterion for this calibration is to ensure the global matching between the performance indices that are directly derived from the raw data and their model-based counterparts in our model-based analysis. We tested whether the results of our model-based analysis change with different adaptation parameters η. The principal results remain unchanged with a high tolerance to the tuning parameter (η in the range from 0.03 to 0.1), validating the robustness of our model.

Our tracking model hypothesizes two likely sequence structures (i.e. level-0 and level-1) and tracks whether these structures could account for the participants' responses. It should be noted that the model has a special structure for level-0 data, where participants' responses are driven by target probability, rather than context- target contingencies.

In addition, our model needs to be initialized before it can track the participant's responses: i.e., a number of initial updates are needed to 'warm-up' the tracking model. We initialized the model by imposing a prior reflecting a possible memory structure used by the participants at the start of training. We initialized the mixture coefficients controlling the degree of a participant's initial preference for one mixture component over the other. Specifically, at level-1 we monitor how the participants switch their prediction strategy from using a level-0 model to using a level-1 model. To reflect the participants' initial preference for a simpler model, we set the initial coefficient of level-0 and level-1 model to w = 0.8 and $1 - w = 0.2$, respectively.

Model-based analysis: context-length model

When fitting the dynamic mixture model (1)–(6) to the response data from a single subject, we obtain time series of mixture coefficients $\{w_t^k\}$ and the context-target contingencies $\{p_t^k\}$, $k = 0, 1$, indexed by the trial index t. The initial settings described above may produce a small systematic variation on the mixture coefficients w_t^k. Therefore, to deal with uncertainty about the participants' initial preferences, we vary the initial mixture coefficients by setting w to a range of 5 values: 0.7, 0.75, 0.8, 0.85, 0.9. For every data level, participant and mixture component k, we then average the resulting 5 trajectories of w_t^k into a single mixture coefficient curve w_t^k. We followed the same method to produce the averaged context-target contingency curves $\{p_t^k\}$.

We now present the analysis methods that enable us to infer characteristics of participant's learning behaviour from the tracking model. We refer to the time series of mixture coefficients, $\{w_t^k\}$, reflecting learning of the memory structure, as the context-length model. In particular, for each participant and data level k_0, we are interested in the dynamics of $\{w_t^{k_0}\}$ across all training blocks for that level k_0 (model order used to generate the data). We denote the series $\{w_t^{k_0}\}$ by $\{\tilde{w}_t\}$. The $\tilde{w}_t$ curves thus represent the dynamics of participants' learning of the correct memory structure when producing predictions for the target items. As the $\tilde{w}_t$ curves have a sigmoid structure in most cases, we represent them through a parameterised sigmoid function

$$\mathcal{W}_t = \mathcal{W}_0 + \frac{\delta_{\mathcal{W}}}{1 + \exp^{-\xi(t-\tau_0)}},$$

where $\mathcal{W}_0$ is the initial mixture coefficient, $\delta_{\mathcal{W}}$ is the difference between the initial and final mixture coefficients, ξ is the learning rate and τ_0 is the transition point. For further analysis, we characterize the $\tilde{w}_t$ curves through the parameter vector $\Theta = (\mathcal{W}_0, \delta_{\mathcal{W}}, \xi, \tau_0)$.

Model-based analysis: predictive contingency model

In addition to learning the context length, the participants need to learn the context-target contingencies. To test whether participants learn context-target contingencies, we used the predictive contingency model (2) from the response tracking model. For each level, we compared this predictive contingency model to two baseline models: (i) the underlying Markov model which is used to generate sequences (i.e. level-1) and (ii) an alternative, less complex model (i.e. an approximate marginalized level-0 model).

We used the expected Kullback-Leiber (KL) divergence from the baseline q to the predictive contingency model p_t (2) to quantify the difference between the two models at trial t:

$$\mathcal{KL}_t = \sum_{\mathcal{C}} p_{\mathcal{C}} \cdot \mathrm{KL}\big(q(s|\mathcal{C})||p_t(s|\mathcal{C})\big).$$

The overall difference between q and p_t is computed as a weighted average of the context-specific differences (KL-divergences) over all contexts in the baseline model. Here, $q(s|\mathcal{C})$ is the context-conditional probability for context $\mathcal{C}$ in q, $p_t(s|\mathcal{C})$ is the participant's predictive contingency for that context at trial t, and $p_{\mathcal{C}}$ is the theoretical distribution of contexts for q obtained from the data generating model. The smaller the divergence between the predictive contingency model and a baseline model, the greater their similarity (in information theoretic terms).

To test which baseline model the participants used, we computed the difference between the expected KL divergence from the less complex model to the predictive contingency model and the expected KL divergence from the underlying Markov model to the predictive contingency model. We refer to this quantity as ΔKL, measuring which of the two baseline models is closer (in the information theoretic sense) to the participant's predictive contingency model. Positive values of ΔKL indicate that participant responses were more aligned with the Markov model used to generate sequences, while negative values indicate responses were based on a less complex model.

Analysis of strategy choice

To estimate the strategy employed by the participants when learning conditional probabilities, we compared individual participant's predictive contingency model (2) to two baseline models: (i) probability matching, where probabilistic distributions are derived from the Markov models that generated the presented sequences (*Model-matching*) and (ii) a probability maximization model, where only the most likely outcomes are allowed for each context (*Model-maximization*). We used Kullback-Leiber (KL) divergence to compare the response distribution to each of these two models. KL is defined as follows:

$$KL = \sum_{c} M(c) log \frac{M(c)}{P(c)}$$

where $P(c)$ and $M(c)$ denote the probability distribution derived from the estimated predictive contingency model and the baseline models (i.e. probability matching or maximization) respectively, across all the probabilistic conditions c (individual targets at level-0 or context-target contingencies at level-1).

We quantified the difference between the KL divergence from *Model-matching* to the human's predictive contingency model and the KL divergence from *Model-maximization* to the predictive contingency model. We refer to this quantity as strategy choice indicated by ΔKL(*Model-maximization*, *Model-matching*). Positive strategy choice values indicate a strategy closer to matching, while negative values indicate a strategy closer to maximization. We computed strategy choice trial-by-trial, resulting in a strategy curve across training for each individual participant. For sequences presented to each participant, we also generated two artificial data sets by simulating responses based on exact matching or maximization.

We then derived an individual learning strategy index by calculating the integral of each participant's strategy curve and subtracting it from the integral of the exact matching curve, as defined by *Model-matching*, across training. We defined the integral curve difference (ICD) between individual strategy and exact matching as the individual strategy index. ICD end refers to ICD calculated over only the last 2 blocks of training. Higher strategy index indicates a strategy closer to maximization.

References

1. Brainard DH (1997) The psychophysics toolbox. *Spatial vision* 10:433-436.

2. Foss A and Zaane OR (2002) A parameterless method for efficiently discovering clusters of arbitrary shape in large datasets. *Data Mining, 2002. ICDM 2003. Proc 2002 IEEE Int Conf,* (IEEE), pp 179-186.

3. Hamerly G and Elkan C (2004) Learning the k in k-means. *Advances in neural information processing systems 16: proceedings of the 2003 conference*, (The MIT Press), p 281.

4. Kass RE and Raftery AE (1995) Bayes factors. *J American Statistical Association* 90(430):773-795.

5. Weinberger KQ and Tesauro G (2007) Metric learning for kernel regression. *International Conference on Artificial Intelligence and Statistics*, 612-19.

References

Arnal, L.H. and A. Giraud (2012). “Cortical oscillations and sensory predictions”. In: *Trends in cognitive sciences* 16.7, pp. 390–398.

Aslin, R.N. and E.L. Newport (2012). “Statistical learning from acquiring specific items to forming general rules”. In: *Current directions in psychological science* 21.3, pp. 170–176.

— (2014). “Distributional language learning: Mechanisms and models of category formation”. In: *Language Learning* 64.s2, pp. 86–105.

Aslin, R.N., J.R. Saffran, and E.L. Newport (1998). “Computation of conditional probability statistics by 8-month-old infants”. In: *Psychological science* 9.4, pp. 321–324.

Baker, R. et al. (2014). “Learning to predict: exposure to temporal sequences facilitates prediction of future events”. In: *Vision research* 99, pp. 124–133.

Baldwin, D. et al. (2008). “Segmenting dynamic human action via statistical structure”. In: *Cognition* 106.3, pp. 1382–1407.

Barnes, R. and M.R. Jones (2000). “Expectancy, attention, and time”. In: *Cognitive psychology* 41.3, pp. 254–311.

Bays, B.C., N.B. Turk-Browne, and A.R. Seitz (2016). “Dissociable behavioural outcomes of visual statistical learning”. In: *Visual Cognition*, pp. 1–26.

Boltz, M.G. (1993). “The generation of temporal and melodic expectancies during musical listening”. In: *Perception & Psychophysics* 53.6, pp. 585–600.

Brady, T.F. and M.M. Chun (2007). “Spatial constraints on learning in visual search: modeling contextual cuing”. In: *Journal of Experimental Psychology: Human Perception and Performance* 33.4, p. 798.

Brady, T.F. and A. Oliva (2008). “Statistical learning using real-world scenes: extracting categorical regularities without conscious intent”. In: *Psychological Science* 19.7, pp. 687–685.

Brainard, D.H. (1997). “The psychophysics toolbox”. In: *Spatial vision* 10, pp. 433–436.

Busch, N.A., J. Dubois, and R. VanRullen (2009). “The phase of ongoing EEG oscillations predicts visual perception”. In: *The Journal of Neuroscience* 29.24, pp. 7869–7876.

Cheke, L.G., J.S. Simons, and N.S. Clayton (2016). "Higher body mass index is associated with episodic memory deficits in young adults". In: *Quarterly Journal of Experimental Psychology*.

Chun, M.M. and Y. Jiang (1998). "Contextual cueing: Implicit learning and memory of visual context guides spatial attention". In: *Cognitive psychology* 36.1, pp. 28–71.

Cleeremans, A. and J.L. McClelland (1991). "Learning the structure of event sequences". In: *Journal of Experimental Psychology: General* 120.3, p. 235.

Cohen, A., R.I. Ivry, and S.W. Keele (1990). "Attention and structure in sequence learning". In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 16.1, p. 17.

Conway, C.M. and M.H. Christiansen (2005). "Modality-Constrained Statistical Learning of Tactile, Visual, and Auditory Sequences". In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 31.1, pp. 24–39.

Correa, A., J. Lupianez, and P. Tudela (2006). "The attentional mechanism of temporal orienting: determinants and attributes". In: *Exp Brain Res* 169, pp. 58–68.

Correa, A., J. Lupiáñez, et al. (2004). "Endogenous temporal orienting of attention in detection and discrimination tasks". In: *Perception & Psychophysics* 66.2, pp. 264–278.

Cravo, A.M. et al. (2013). "Temporal expectation enhances contrast sensitivity by phase entrainment of low-frequency oscillations in visual cortex". In: *The Journal of Neuroscience* 33.9, pp. 4002–4010.

Dale, R. and M.H. Christiansen (2004). "Active and passive statistical learning: Exploring the role of feedback in artificial grammar learning and language". In: *Proceedings of the 26th Annual Conference of the Cognitive Science Society*, pp. 262–267.

Derks, P.L. and M.I. Paclisanu (1967). "Simple Strategies in Binary Prediction by Children and Adults". In: *Journal of Experimental Psychology* 73.2, p. 278.

Doherty, J.R. et al. (2005). "Synergistic effect of combined temporal and spatial expectations on visual attention". In: *The Journal of neuroscience* 25.36, pp. 8259–8266.

Durrant, S.J., S.A. Cairney, and P.A. Lewis (2013). "Overnight consolidation aids the transfer of statistical knowledge from the medial temporal lobe to the striatum". In: *Cerebral Cortex* 23, pp. 2467–2478.

Durrant, S.J., C. Taylor, et al. (2011). "Sleep-dependent consolidation of statistical learning". In: *Neuropsychologia* 49.5, pp. 1322–1331.

Edwards, J.D. et al. (2006). "The useful field of view test: normative data for older adults". In: *Archives of Clinical Neuropsychology* 21.4, pp. 275–286.

Fernandes, T., R. Kolinsky, and P. Ventura (2010). "The impact of attention load on the use of statistical information and coarticulation as speech segmentation cues". In: *Attention, Perception and Psychophysics* 72.6, pp. 1522–1532.

Fiser, J. and R.N. Aslin (2001). "Unsupervised statistical learning of higher-order spatial structures from visual scenes". In: *Psychological science* 12.6, pp. 499–504.

— (2002a). "Statistical learning of higher-order temporal structure from visual shape sequences". In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 28.3, p. 458.

— (2002b). "Statistical learning of new visual feature combinations by infants". In: *Proceedings of the National Academy of Sciences* 99.24, pp. 15822–15826.

— (2005). "Encoding multielement scenes: statistical learning of visual feature hierarchies". In: *Journal of Experimental Psychology: General* 134.4, p. 521.

Frost, R. et al. (2015). "Domain generality versus modality specificity: the paradox of statistical learning". In: *Trends in Cognitive Sciences* 19.3.

Gao, J. and J.E. Corter (2015). "Striving for perfection and falling short: The influence of goals on probability matching". In: *Memory and cognition* 43.5, pp. 748–759.

Gerken, L. (2006). "Decisions, decisions: Infant language learning when multiple generalizations are possible". In: *Cognition* 98.3, B67–B74.

— (2010). "Infants use rational decision criteria for choosing among models of their input". In: *Cognition* 115.2, pp. 362–366.

Goldwater, S., T.L. Griffiths, and M. Johnson (2009). "A Bayesian framework for word segmentation: Exploring the effects of context". In: *Cognition* 112, pp. 21–54.

Gómez, R.L. (1997). "Transfer and Complexity in Artificial Grammar Learning". In: *Cognitive Psychology* 33, pp. 154–207.

Guo, K. et al. (2004). "Effects on orientation perception of manipulating the spatio-temporal prior probability of stimuli". In: *Vision research* 44.20, pp. 2349–2358.

Howard, J.H. et al. (2008). "Implicit learning of predictive relationships in three-element visual sequences by young and old adults". In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 34.5, p. 1139.

Hunt, R.H. and R.N. Aslin (2001). "Statistical learning in a serial reaction time task: access to separable statistical cues by individual learners". In: *Journal of Experimental Psychology: General* 130.4, p. 658.

Jones, M.R. and M. Boltz (1989). "Dynamic attending and responses to time". In: *Psychological review* 96.3, p. 459.

Jones, M.R., H. Moynihan, et al. (2002). "Temporal aspects of stimulus-driven attending in dynamic arrays". In: *Psychol Sci* 13.4, pp. 313–319.

Karuza, E.A. et al. (2013). “The neural correlates of statistical learning in a word segmentation task: An fMRI study”. In: *Brain and language* 127.1, pp. 46–54.

Kim, R. et al. (2009). “Testing assumptions of statistical learning: is it long-term and implicit?” In: *Neuroscience letters* 461.2, pp. 145–149.

Kirkham, N.Z., J.A. Slemmer, and S.P. Johnson (2002). “Visual statistical learning in infancy: Evidence for a domain general learning mechanism”. In: *Cognition* 83.2, B35–B42.

Kluger, A.N. and A. DeNisi (1996). “The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory”. In: *Psychological bulletin* 119.2, p. 254.

Koehler, D.J. and G. James (2010). “Probability matching and strategy availability”. In: *Memory and Cognition* 38.6, pp. 667–676.

— (2014). “Probability matching, fast and slow”. In: *The Psychology of Learning and Motivation* 61, pp. 103–132.

Lange, K. (2010). “Can a regular context induce temporal orienting to a target sound?” In: *International Journal of Psychophysiology* 78.3, pp. 231–238.

Large, E.W. and M.R. Jones (1999). “The dynamics of attending: How people track time-varying events”. In: *Psychological review* 106.1, p. 119.

Lasley, D.J. and T. Cohn (1981). “Detection of a luminance increment: effect of temporal uncertainty”. In: *JOSA* 71.7, pp. 845–850.

Lotz, A. and A. Kinder (2006). “Transfer in artificial grammar learning: The role of repetition information”. In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 32.4, p. 707.

Luck, S.J. and E.K. Vogel (1997). “The capacity of visual working memory for features and conjunctions”. In: *Nature* 390.6657, pp. 279–281.

Marcus, G.F., K.J. Fernandes, and S.P. Johnson (2007). “Infant Rule Learning Facilitated by Speech”. In: *Psychological Science* 18.5, pp. 387–391.

Marcus, G.F., S. Vijayan, et al. (1999). “Rule Learning by Seven-Month-Old Infants”. In: *Science* 283, pp. 77–80.

Maye, J., J.F. Werker, and L. Gerken (2002). “Infant sensitivity to distributional information can affect phonetic discrimination”. In: *Cognition* 82.3, B101–B111.

McMahon, A.J. and M.H. Scheel (2010). “Glucose promotes controlled processing: Matching, maximizing, and root beer”. In: *Judgment and Decision Making* 5.6, p. 450.

Misyak, J.B., M.H. Christiansen, and J. Bruce Tomblin (2010). “Sequential Expectations: The Role of Prediction-Based Learning in Language”. In: *Topics in Cognitive Science* 2.1, pp. 138–153.

Murphy, F.C. et al. (2003). "Neuropsychological impairment in patients with major depressive disorder: the effects of feedback on task performance". In: *Psychological medicine* 33.03, pp. 455–467.

Newell, B.R., D.J. Koehler, et al. (2013). "Probability matching in risky choice: The interplay of feedback and strategy availability". In: *Memory and cognition* 41.3, pp. 329–338.

Newell, B.R. and T. Rakow (2007). "The role of experience in decisions from description". In: *Psychonomic Bulletin and Review* 14.6, pp. 1133–1139.

Niemi, P. and R. Näätänen (1981). "Foreperiod and simple reaction time." In: *Psychological Bulletin* 89.1, p. 133.

Nissen, M. and P. Bullemer (1987). "Attentional requirements of learning: Evidence from performance measures". In: *Cognitive psychology* 19.1, pp. 1–32.

Nobre, A.C., A. Correa, and J.T. Coull (2007). "The hazards of time". In: *Current opinion in neurobiology* 17.4, pp. 465–470.

Olson, I.R. and M.M. Chun (2001). "Temporal contextual cuing of visual attention". In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 27.5, p. 1299.

Otsuka, S. and J. Saiki (2016). "Gift from statistical learning: Visual statistical learning enhances memory for sequence elements and impairs memory for items that disrupt regularities". In: *Cognition* 147, pp. 113–126.

Pelli, D.G. (1997). "The VideoToolbox software for visual psychophysics: Transforming numbers into movies". In: *Spatial vision* 10.4, pp. 437–442.

Perruchet, P. and S. Pacton (2006). "Implicit learning and statistical learning: one phenomenon, two approaches". In: *TRENDS in Cognitive Sciences* 10.5.

Potter, M.C. et al. (2002). "Recognition memory for briefly presented pictures: the time course of rapid forgetting". In: *Journal of Experimental Psychology: Human Perception and Performance* 28.5.

Praamstra, P. et al. (2006). "Neurophysiology of implicit timing in serial choice reaction-time performance". In: *The Journal of neuroscience* 26.20, pp. 5448–5455.

Reber, A.S. (1967). "Implicit learning of artificial grammars". In: *Journal of verbal learning and verbal behavior* 6.6, pp. 855–863.

Reeder, P.A., E.L. Newport, and R.N. Aslin (2009). "The role of distributional information in linguistic category formation". In: *Proceedings of the 31st annual meeting of the Cognitive Science Society*, pp. 2564–2569.

— (2010). "Novel words in novel contexts: The role of distributional information in form-class category learning". In: *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*, pp. 2063–2068.

Rohenkohl, G. and A.C. Nobre (2011). “Alpha oscillations related to anticipatory attention follow temporal expectations”. In: *The Journal of Neuroscience* 31.40, pp. 14076–14084.

Rolke, B. and P. Hofmann (2007). “Temporal uncertainty degrades perceptual processing”. In: *Psychonomic Bulletin & Review* 14.3, pp. 522–526.

Rowland, L.A. and D.R. Shanks (2006). “Attention modulates the learning of multiple contingencies”. In: *Psychometric Bulletin and Review* 13.4, pp. 643–648.

Saffran, J.R. (2001). “Words in a sea of sounds: The output of infant statistical learning”. In: *Cognition* 81.2, pp. 149–169.

— (2002). “Constraints on statistical language learning”. In: *Journal of Memory and Language* 47.1, pp. 172–196.

Saffran, J.R., R.N. Aslin, and E.L. Newport (1996). “Statistical learning by 8-month-old infants”. In: *Science* 274.5294, pp. 1926–1928.

Saffran, J.R., E.K. Johnson, et al. (1999). “Statistical learning of tone sequences by human infants and adults”. In: *Cognition* 70.1, pp. 27–52.

Saffran, J.R., E.L. Newport, and R.N. Aslin (1996). “Word segmentation: The role of distributional cues”. In: *Journal of memory and language* 35.4, pp. 606–621.

Saffran, J.R., E.L. Newport, R.N. Aslin, et al. (1997). “Incidental language learning: Listening (and learning) out of the corner of your ear”. In: *Psychological science* 8.2, pp. 101–105.

Saffran, J.R. and E.D. Thiessen (2003). “Pattern induction by infant language learners”. In: *Developmental psychology* 39.3, p. 484.

Schapiro, A.C., E. Gregory, et al. (2014). “The necessity of the medial temporal lobe for statistical learning”. In: *Journal of cognitive neuroscience* 26.8, pp. 1736–1747.

Schapiro, A.C., L.V. Kustner, and N.B. Turk-Browne (2012). “Shaping of object representations in the human medial temporal lobe based on temporal regularities”. In: *Current Biology* 22.17, pp. 1622–1627.

Schulze, C., D. van Ravenzwaaij, and B.R. Newell (2015). “Of matchers and maximizers: How competition shapes choice under risk and uncertainty”. In: *Cognitive psychology* 78, pp. 78–98.

Schwarb, H. and E.H. Schumacher (2012). “Generalized lessons about sequence learning from the study of the serial reaction time task”. In: *Advances in cognitive psychology* 8.2, pp. 165–178.

Schwartze, M. et al. (2011). “Temporal regularity effects on pre-attentive and attentive processing of deviance”. In: *Biological psychology* 87.1, pp. 146–151.

Shanks, D.R., R.J. Tunney, and J.D. McCarthy (2002). “A re-examination of probability matching and rational choice”. In: *Journal of Behavioral Decision Making* 15.3, pp. 233–250.

Toro, J.M., S. Sinnett, and S. Soto-Faraco (2005). "Speech segmentation by statistical learning depends on attention". In: *Cognition* 97, B25–B34.

Turk-Browne, N.B., P.J. Isola, et al. (2008). "Multidimensional Visual Statistical Learning". In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 34.2, pp. 399–407.

Turk-Browne, N.B., J.A. Jungé, and B.J. Scholl (2005). "The Automaticity of Visual Statistical Learning". In: *Journal of Experimental Psychology: General* 134.4, pp. 552–564.

Turk-Browne, N.B. and B.J. Scholl (2009). "Flexible Visual Statistical Learning: Transfer Across Space and Time". In: *Human Perception and Performance* 35.1, pp. 195–202.

Turk-Browne, N.B., B.J. Scholl, et al. (2010). "Implicit perceptual anticipation triggered by statistical learning". In: *The Journal of Neuroscience* 30.33, pp. 11177–11187.

Vokey, J.R. and L.R. Brooks (1992). "Salience of item knowledge in learning artificial grammars". In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 18.2, p. 328.

Vulkan, N. (2000). "An economist's perspective on probability matching". In: *Journal of economic surveys* 14.1, pp. 101–118.

Wang, R. et al. (in review). "Extracting behaviorally relevant statistics promotes structural- over rote- learning".

West, R.F. and K.E. Stanovich (2003). "Is probability matching smart? Associations between probabilistic choices and cognitive ability". In: *Memory & Cognition* 31.2, pp. 243–251.

Westheimer, G. and E. Ley (1996). "Temporal uncertainty effects on orientation discrimination and stereoscopic thresholds". In: *JOSA A* 13.4, pp. 884–886.

Wolford, G. et al. (2004). "Searching for patterns in random sequences". In: *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale* 58.4, p. 221.

Woodrow, H. (1914). "The measurement of attention". In: *The Psychological Monographs* 17.5, p. i.

Zanto, T.P., J.S. Snyder, and E.W. Large (2006). "Neural correlates of rhythmic expectancy". In: *Advances in Cognitive Psychology* 2.2–3, pp. 221–231.